

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:16:04 ; Search time 27.947 Seconds
(without alignments)
1257.748 Million cell updates/sec

Title: US-10-801-517-2
Perfect score: 412
Sequence: 1 SDVYCVCFVLKVEVTKLID.....ILLEEVSPBLVCSMLHLCSSG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412	100.0	80	2	AAR70784 Saposin-C
2	412	100.0	80	2	AAR70784 Human sap
3	412	100.0	80	4	AAR70784 Human sap
4	412	100.0	80	7	ABU05697 Human sap
5	412	100.0	80	8	ABU05697 Spingoli
6	412	100.0	80	8	ADQ94329 Human Pre
7	412	100.0	80	8	ADQ94329 Human Pre
8	412	100.0	80	9	ADQ94329 Human Pre
9	412	100.0	80	9	ADQ94329 Human Pre
10	412	100.0	210	6	ABU05201 Human exp
11	412	100.0	385	6	ABR41750 Human pro
12	412	100.0	479	6	ABR39442 Human GEN
13	412	100.0	522	8	ADU24090 Human cys
14	412	100.0	523	4	ABR31916 Amino aci
15	412	100.0	523	6	ABU05211 Human exp
16	412	100.0	524	2	AAR70783 Prosaposi
17	412	100.0	524	3	AAW85652 Human pro
18	412	100.0	524	3	AAW85652 Human pro
19	412	100.0	524	6	ABU79099 Lip-TAA b
20	412	100.0	524	6	ABU05200 Human exp
21	412	100.0	524	6	ABU05207 Human exp
22	412	100.0	524	6	ABU05203 Human exp
23	412	100.0	524	6	ABU07340 Human exp
24	412	100.0	524	6	ABU05216 Human exp

BEST AVAILABLE COPY

BEST AVAILABLE COPY

25	412	100.0	524	6	ABU05208 Human exp
26	412	100.0	524	6	ABU05214 Human exp
27	412	100.0	524	6	ABU05215 Human exp
28	412	100.0	524	6	ABU05199 Human exp
29	412	100.0	524	6	ABU05212 Human exp
30	412	100.0	524	6	ABU05213 Human exp
31	412	100.0	524	6	ABU05205 Human exp
32	412	100.0	524	7	ADU43340 Superanti
33	412	100.0	524	7	ADJ69401 Human hea
34	412	100.0	524	8	ADQ94328 Human Pre
35	412	100.0	524	8	ADQ94328 Human Pre
36	412	100.0	524	8	ABM81149 Tumour-as
37	412	100.0	524	8	ADU87894 Human pro
38	412	100.0	524	8	ADU48630 Human pro
39	412	100.0	524	9	ADW80727 Human pro
40	412	100.0	524	9	ADX06774 Cyclin-de
41	412	100.0	524	9	ADY14302 PRO poly
42	412	100.0	526	6	ABU05209 Human exp
43	412	100.0	527	4	ABP31915 Amino aci
44	412	100.0	527	5	ABP68602 Human pan
45	412	100.0	527	6	ABU79100 Lip-TAA b

ALIGNMENTS

RESULT 1
AAR70784
ID AAR70784 standard; protein; 80 AA.

XX AAR70784;

XX AC

DT 25-MAR-2003 (revised)

DT 30-AUG-1995 (first entry)

XX Saposin-C.

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
 CC correct PI field.)
 XX
 SQ Sequence 80 AA;
 Query Match 100.0%; Score 412; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTYSSILS 60
 DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTYSSILS 60
 QY 61 ILLEEVSPELVCSMLHLCSG 80
 DB 61 ILLEEVSPELVCSMLHLCSG 80
 RESULT 2
 ID AAW85653 standard; peptide; 80 AA.
 XX
 AC AAW85653;
 XX
 DT 19-JUL-1999 (first entry)
 XX
 DE Human saposin C.
 XX
 KW Prosaposin, saposin; prosaptides; prosaposin receptor agonists; PRA;
 KW peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
 KW therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;
 KW cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
 KW Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;
 KW congestive heart failure; multiple sclerosis;
 KW acute disseminated inflammatory leukoencephalitis;
 KW progressive multifocal leukoencephalitis; Alzheimer's disease;
 KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
 KW ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
 KW cerebral malaria; HTLV; neuropathy;
 KW inflammatory neurodegenerative disease; toxin-induced liver disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9912559-A1.
 XX
 PD 18-MAR-1999.
 XX
 PF 09-SEP-1998; 98WO-US019216.
 XX
 PR 09-SEP-1997; 97US-0058352P.
 PR 04-JUN-1998; 98US-0088129P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI O'Brien JS;
 XX
 DR WPI; 1999-229139/19.
 XX
 XX Use of prosaposin receptor agonist.
 XX
 PS Claim 7; Fig 2; 90pp; English.
 XX
 CC Prosaposin is a 70kDa glycoprotein which is proteolytically processed to
 CC generate saposins A, B, C and D, all of which are similar to each other
 CC and have a similar placement of six cysteines, a glycosylation site and
 CC conserved proline residues. Prosaposin, saposin C and prosaposin derived
 CC peptides (prosaptides), have therapeutic applications in promoting
 CC recovery after toxic, traumatic, myocardial ischemic, degenerative and
 CC inherited lesions to the peripheral and central nervous system.
 CC Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-
 CC induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt
 CC dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2,
 CC releasing Bcl-2 and its family members which inhibit caspases, thereby

CC inhibiting apoptosis. An additional mechanism whereby PRAs inhibit
 CC apoptosis is by blocking activation of JNK, a proapoptotic signaling
 CC component. Within several minutes after binding to the receptor, PRAs
 CC block JNK activation induced by tumor necrosis factor-alpha (TNF alpha).
 CC The activation of JNK by TNF alpha is another well known mechanism for
 CC TNF alpha-induced, as well as other proinflammatory cytokine-induced
 CC apoptosis. The method can be used for inhibiting apoptosis which is
 CC caspase-mediated or induced by a proinflammatory cytokine, for example
 CC TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis
 CC associated with a disorder such as e.g. rheumatoid arthritis, Crohn's
 CC disease, irritable bowel syndrome, asthma, cardiac infarction, congestive
 CC heart failure, multiple sclerosis, acute disseminated inflammatory
 CC leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's
 CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
 CC disease, ischemic heart disease, Guillain-Barre disease, traumatic brain
 CC injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral
 CC malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and
 CC toxin-induced liver disease. Saposin C acts as a prosaposin receptor
 CC agonist
 XX
 SQ Sequence 80 AA;
 Query Match 100.0%; Score 412; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTYSSILS 60
 DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTYSSILS 60
 QY 61 ILLEEVSPELVCSMLHLCSG 80
 DB 61 ILLEEVSPELVCSMLHLCSG 80
 RESULT 3
 ID AAW05697 standard; protein; 80 AA.
 XX
 AC AAW05697;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human Saponin C, SapC.
 XX
 KW Human; glucocerebrosidase; GCB; lysosomal storage disease;
 KW Gaucher's disease; Fabry's disease; Farber's disease;
 KW G_m_1 gangliosidosis; Tay-Sach's disease; Niemann-Pick disease;
 KW Shindler disease; Hunter syndrome; Sly syndrome; Huler syndrome;
 KW Scheie syndrome; Saponin C; SapC.
 XX
 OS Homo sapiens.
 XX
 PN WO200149830-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-DK000743.
 XX
 PR 30-DEC-1999; 99DK-00001891.
 PR 02-JUN-2000; 2000DK-0000865.
 PR 02-JUN-2000; 2000DK-0000866.
 PR 30-JUN-2000; 2000DK-00001027.
 XX
 PA (MAXY-) MAXYGEN APS.
 XX
 PI Okkels JS, Jensen AD, Halkier T, Jensen RB, Schambye HT;
 DR WPI; 2001-465259/50.
 XX
 PT Improved lysosomal enzymes and lysosomal enzyme activators useful for
 PT treating Gaucher's disease.
 XX

Example 5; Page 96; 97pp; English.

The sequence represents human Saponin C (SapC), an essential co-factor for the lysosomal enzyme glucocerebrosidase. GCB. GCB is the enzyme involved in Gaucher's disease, a lysosomal storage disease. The invention relates to introducing new glycosylation sites into lysosomal enzymes/activators like GCB to improve their bioactivity. The novel polypeptides are used for the prevention and treatment of Gaucher's disease, Fabry's disease, Farber's disease, G.M.I gangliosidosis, Tay-Sach's disease, Niemann-Pick disease, Shindler disease, Hunter syndrome, Sly syndrome, Hurler and Huler/Scheie syndromes

Sequence 80 AA;

Query Match 100.0%; Score 412; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPSLSECOVVDVTYSSILS 60
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPSLSECOVVDVTYSSILS 60
Qy 61 ILLEVSPELVCSMLHLCSSG 80
Db 61 ILLEVSPELVCSMLHLCSSG 80

RESULT 4
ABU6252
ID ABU62252 standard; protein; 80 AA.
XX AC ABU62252;
XX XX
DT 29-AUG-2003 (first entry)
XX DE
XX Spingolipid activator protein C (saposin C).

XX Sphingolipid activator protein; saposin; neuroprotective; human;
KW Sphingolipid activator phospholipid; fusogenic protein; prosaposin;
KW gene therapy; anionic phospholipid; fusogenic protein; prosaposin;
KW Gauchers disease; saposin C.
XX XX
OS Homo sapiens.
XX XX
XX US2003095999-A1.
XX XX
XX 22-MAY-2003.
XX XX
XX 09-FEB-2001; 2001US-00780438.
XX XX
XX 11-FEB-2000; 2000US-0181754P.
XX XX
XX (Q1XX/) Q1 X.
XX XX
XX Q1 X;
XX XX
XX WPI; 2003-512933/48.
XX XX

Delivering pharmaceutical agent through membrane used for treating Gauchers disease, by applying composition comprising anionic phospholipids and fusogenic protein derived from prosaposin in carrier to membrane.

Disclosure; Fig 4; 19pp; English.

The invention describes a method of delivering a pharmaceutical agent through a membrane comprising applying to the membrane a composition (I) comprising: (1) anionic phospholipids; (2) a pharmaceutical agent contained within the phospholipids; and (3) a fusogenic protein or polypeptide derived from prosaposin in a carrier. The method can be used for delivering pharmaceutical agents through a biological membrane in cosmetic and medicinal applications, particularly for treating Gauchers disease. This is the amino acid sequence of sphingolipid activator protein C (saposin C) for use in the delivery composition

XX SQ Sequence 80 AA;

Query Match 100.0%; Score 412; DB 7; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPSLSECOVVDVTYSSILS 60
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPSLSECOVVDVTYSSILS 60
Qy 61 ILLEVSPELVCSMLHLCSSG 80
Db 61 ILLEVSPELVCSMLHLCSSG 80

RESULT 5
ADQ94329
ID ADQ94329 standard; protein; 80 AA.
XX AC ADQ94329;
XX XX
DT 23-SEP-2004 (first entry)
XX DE
XX Human Prosaposin protein, C-domain.

XX Human; prosaposin; C-domain; prosaposin receptor agonist;
KW neuropathic pain; neurite outgrowth; neural cell death; myelination;
KW demyelination; neuropathy; peripheral nerve disorder; neuroma;
KW nerve compression; nerve crush; nerve stretch;
KW incomplete nerve transection; mononeuropathy; polyneuropathy.
XX XX
OS Homo sapiens.
XX XX
XX US2004121958-A1.
XX XX
XX 24-JUN-2004.
XX XX
XX 24-DEC-2003; 2003US-00746442.
XX XX
XX 05-MAR-1996; 96US-00611307.
XX PR 05-MAR-1997; 97WO-US004143.
XX PR 11-SEP-1997; 97US-00928074.
XX XX
XX (REGC) UNIV CALIFORNIA.
XX XX
XX O'brien JS;
XX XX
XX WPI; 2004-468194/44.
XX XX
XX New prosaposin receptor agonist, useful in alleviating or treating neuropathic pain, inhibiting the onset of neuropathic pain, neural cell death, demyelination, or sensory or motor neuropathy, and stimulating neurite outgrowth.

Disclosure; SEQ ID NO 24; 33pp; English.

The invention relates to a prosaposin receptor agonist comprising a defined amino acid sequence of 14-50 amino acids conforming to the consensus sequence appearing as ADQ94330. The agonists are based on peptides derived from the C domain of human prosaposin (or homologous peptides from other proteins). Also included are a pharmaceutical composition comprising the prosaposin receptor agonist in a pharmaceutical carrier, a method of alleviating neuropathic pain in a subject, a method of inhibiting the onset of neuropathic pain in a cell death, promoting myelination or inhibiting demyelination) and a method of inhibiting sensory or motor neuropathy. The neuropathic pain results from a peripheral nerve disorder, e.g. neuroma, nerve compression, nerve crush, nerve stretch and incomplete nerve transection, mononeuropathy or polyneuropathy or results from a disorder of dorsal root ganglia, spinal cord, brainstem, thalamus, or cortex. The prosaposin receptor agonist, composition, and methods are useful in

Wed Jan 18 11:58:41 2006

CC The subject is a mammal, preferably a human. The cell death preferably
 CC occurs through apoptosis. The agent is useful in treating cancer by
 CC inducing cell death in hyperproliferating cells. The present sequence
 CC represents a human Saposin C polypeptide.

XX Sequence 80 AA;
 SQ Query Match 100.0%; Score 412; DB 8; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSECEQVVDVTYSSILS 60
 DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSECEQVVDVTYSSILS 60
 QY 61 ILLEEVSPELVCSMLHLCSSG 80
 DB 61 ILLEEVSPELVCSMLHLCSSG 80

RESULT 7
 ADW80728
 ID ADW80728 standard; protein; 80 AA.

XX ADW80728;
 AC ADW80728;
 XX 07-APR-2005 (first entry)
 DT Human Saposin C, SEQ ID NO:24.
 XX Pharmaceutical; neuropathy; analgesic; antidiabetic; cell death;
 XX diabetes; neurological disease; neuropathic pain; saposin C.
 XX Homo sapiens.
 OS US6849602-B1.
 XX 01-FEB-2005.
 PD 11-SEP-1997; 97US-00928074.
 PF 05-MAR-1996; 96US-00611307.
 PR 05-MAR-1997; 97WO-US004143.
 XX (REGC) UNIV CALIFORNIA.
 PA Obrien JS;
 PI WPI; 2005-129968/14.
 XX New pharmaceutical composition comprising a proapoptin receptor agonist
 PT and a carrier, where the composition is in a controlled release,
 PT liposomal or unit dosage form, useful for alleviating neuropathic pain
 PT resulting from diabetes.
 PS Disclosure; SEQ ID NO 24; 35pp; English.

XX The invention relates to pharmaceutical composition comprising a
 CC proapoptin receptor agonist and a carrier, where the composition is in a
 CC controlled release, liposomal or unit dosage form. The invention also
 CC relates to proapoptin receptor agonists and the use of these agonists for
 CC stimulating neurite outgrowth, inhibiting neural cell death, promoting
 CC myelination, inhibiting neural demyelination and inhibiting sensory or
 CC motor neuropathy. The pharmaceutical composition is useful for
 CC alleviating neuropathic pain resulting from diabetes. The present
 CC sequence is human saposin C, which is derived from proapoptin. Note: The
 CC present sequence is given in the sequence listing, but is not further
 CC referred to in the specification.

XX Sequence 80 AA;
 SQ Query Match 100.0%; Score 412; DB 9; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;

CC alleviating or treating neuropathic pain, inhibiting the onset of
 CC neuropathic pain, neural cell death, demyelination, or sensory or motor
 CC neuropathy, and in stimulating neurite outgrowth. The present sequence
 CC represents the C-domain of Human proapoptin.

XX Sequence 80 AA;
 SQ Query Match 100.0%; Score 412; DB 8; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSECEQVVDVTYSSILS 60
 DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSECEQVVDVTYSSILS 60
 QY 61 ILLEEVSPELVCSMLHLCSSG 80
 DB 61 ILLEEVSPELVCSMLHLCSSG 80

RESULT 6
 ADU48631
 ID ADU48631 standard; protein; 80 AA.

XX ADU48631;
 AC ADU48631;
 XX 27-JAN-2005 (first entry)
 DT Human Saposin C polypeptide.
 XX Proapoptin; saposin C; phosphatidylserine; dioleoylphosphatidylserine;
 XX cytosolic; apoptosis stimulator; human.
 OS Homo sapiens.
 OS WO2004096159-A2.
 XX 11-NOV-2004.
 PD 17-MAR-2004; 2004WO-US008020.
 PF 28-APR-2003; 2003US-0466166P.
 PR 16-MAR-2004; 2004US-00801517.
 XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
 PA Qi X;
 PI WPI; 2004-804606/79.

XX New agent comprising an inner leaflet component (preferably
 PT phosphatidylserine) and a proapoptin-related polypeptide with plasma
 PT membrane affinity comprising a sequence of 524 or 80 amino acids, useful
 PT in treating cancer.

XX Claim 1; SEQ ID NO 2; 53pp; English.
 PS The invention relates to a agent comprising an inner leaflet component
 CC and a proapoptin-related polypeptide. The proapoptin-related polypeptide
 CC is selected from ADU48630 or ADU48631, where the polypeptide retains
 CC plasma-membrane affinity. The inner leaflet component is
 CC phosphatidylserine or its structural analogue, and is especially
 CC dioleoylphosphatidylserine. The agent promotes cell death in
 CC hyperproliferating cells (tumour cells and cancer cells). The agent can
 CC be used in a method (M1) for modulating the distribution of an inner
 CC leaflet component in a subject; treating (M3) cancer in a subject. In M1
 CC (M2) tumour volume in a subject; treating (M3) cancer in a subject. The
 CC the concentration of the inner leaflet component in the outer leaflet of
 CC the plasma membrane is altered (preferably increased). The distribution
 CC of the inner leaflet component is modulated in hyper-proliferating cells
 CC (tumour cells and cancer cells). In M2 and M3 the agent promotes cell
 CC death in hyperproliferating cells which are preferably selected from
 CC sarcoma, neuroblastoma, breast carcinoma, and squamous carcinoma cells.

us-10-801-517-2.rag

Wed Jan 18 11:58:41 2006

CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 210 AA;
 Query Match 100.0%; Score 412; DB 6; Length 210;
 Best Local Similarity 100.0%; Pred. No. 8.6e-40; Indels 0; Gaps 0;
 Matches 80; Conservative 0; Mismatches 0;
 QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDYTGSSILS 60
 DB 8 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDYTGSSILS 67
 QY 61 ILLEEVSPELVCSMLHLCSG 80
 DB 68 ILLEEVSPELVCSMLHLCSG 87
 RESULT 10
 ABR41750
 ID ABR41750 standard; protein; 385 AA.
 AC ABR41750;
 DT 02-JUN-2003 (first entry)
 XX Human DITHP biochemical pathway protein.
 XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW biochemical pathway.
 XX Homo sapiens.
 XX WO200297031-A2.
 XX 05-DEC-2002.
 XX 27-MAR-2002; 2002WO-US010056.
 XX 28-MAR-2001; 2001US-0279619P.
 XX 29-MAR-2001; 2001US-0280067P.
 XX 29-MAR-2001; 2001US-0280068P.
 XX 16-MAY-2001; 2001US-0291280P.
 XX 17-MAY-2001; 2001US-0291829P.
 XX 17-MAY-2001; 2001US-0281849P.
 XX 19-JUN-2001; 2001US-0299428P.
 XX 20-JUN-2001; 2001US-0299776P.
 XX 20-JUN-2001; 2001US-0300001P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 XX Dufour GE, Hillman JL, Yu JY, Tuason D, Amshey SR;
 XX Daughtery SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;
 XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX WPI; 2003-129518/12.

DR N-PSDB; ACC46687.
 XX Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
 XX Claim 27; SEQ ID NO 1285; 591pp; English.
 XX The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
 CC polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
 CC detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; gastrointestinal disorders; transport
 CC disorders; neurological disorders; gastrointestinal disorders; they may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a DITHP protein which is involved in a
 CC biochemical pathway. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 385 AA;
 Query Match 100.0%; Score 412; DB 6; Length 385;
 Best Local Similarity 100.0%; Pred. No. 1.9e-39; Indels 0; Gaps 0;
 Matches 80; Conservative 0; Mismatches 0;
 QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDYTGSSILS 60
 DB 172 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDYTGSSILS 231
 QY 61 ILLEEVSPELVCSMLHLCSG 80
 DB 232 ILLEEVSPELVCSMLHLCSG 251
 RESULT 11
 ABR39442
 ID ABR39442 standard; protein; 479 AA.
 AC ABR39442;
 XX 12-JUN-2003 (first entry)
 XX Human GENSET polypeptide clone name SAP-MU-10.
 XX GENSET; cytostatic; gene therapy; cancer; transgenic; human.
 XX Homo sapiens.
 XX WO2003014151-A2.
 XX 20-FEB-2003.
 XX 15-OCT-2001; 2001WO-IB002321.
 XX

PR 10-AUG-2001; 2001US-0311305P.
 PR 24-AUG-2001; 2001US-0314734P.
 PR 07-SEP-2001; 2001US-0318204P.
 PR 01-OCT-2001; 2001US-0326470P.
 XX (GEST) GENSET SA.
 PA PA
 PI Bejamin S, Tanaka H;
 XX XX
 DR WPI; 2003-256539/25.
 DR N-PSDB; ABZ76265.
 XX XX
 PT New GENSET gene, useful for preparing a composition for treating GENSET-
 PT related disorders.
 XX XX
 PS Claim 2; Page 288-289; 301pp; English.
 XX XX
 CC The invention relates to isolated GENSET polynucleotides and encoded
 CC polypeptides. The GENSET gene is useful for preparing a composition for
 CC treating GENSET-related disorders e.g., cancer. Sequences ABR39423-448
 CC represent the novel GENSET polypeptide sequences
 XX XX
 SQ Sequence 479 AA;
 Query Match 100.0%; Score 412; DB 6; Length 479;
 Best Local Similarity 100.0%; Pred. No. 2.5e-39;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SDVYCEVCEFLVKVETKLDNNKTEKILDAFKMCKLPSLSSECQVVDYTGSSILS 60
 DB 314 SDVYCEVCEFLVKVETKLDNNKTEKILDAFKMCKLPSLSSECQVVDYTGSSILS 373
 QY 61 ILLEEVSPVLVCSMLHLCSG 80
 DB 374 ILLEEVSPVLVCSMLHLCSG 393
 RESULT 12
 ADU24090
 ID ADU24090 standard; protein; 522 AA.
 XX XX
 AC ADU24090;
 XX XX
 DT 27-JAN-2005 (first entry)
 XX XX
 DE Human cysteine peptidase enzyme #12.
 XX XX
 KW Protease; cancer; immune-related disorder; cardiovascular disease;
 KW brain-associated disease; neuronal-associated disease;
 KW metabolic disorder; haematopoietic; cytostatic; immunosuppressive;
 KW cardiovascular; cerebroprotective; neuroprotective; human;
 KW cysteine peptidase; enzyme; peripheral nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; infection; ocular disease; migraine; pain;
 KW sexual dysfunction; mood disorder; attention disorder; cognition disorder;
 KW hypertension; hypotension; psychotic disorder; neurological disorder;
 KW dyskinesia; organ transplant rejection.
 XX OS
 OS Homo sapiens.
 XX XX
 FH Key Location/Qualifiers
 FT Domain 258..444
 FT /note = Protease active domain
 XX XX
 US2004219609-A1.
 XX XX
 PD 04-NOV-2004.
 XX XX
 PF 11-JUL-2003; 2003US-00618281.
 XX XX
 PR 12-JUL-2002; 2002US-0395325P.
 XX XX
 PA (DAYA/) DAY A G.
 (ESTE/) ESTELL D A.
 (LYON/) LYONS E H.
 (YAO/) YAO J.
 Day AG, Estell DA, Lyons EH, Yao J;
 WPI; 2004-794441/78.
 Identifying compound that modulates activity of protease, by contacting
 protease with test compound, measuring activity of protease before and
 after contacting step, and determining whether test compound modulates
 activity of protease.
 Claim 1; SEQ ID NO 38; 206pp; English.
 The invention relates to a method for identifying a compound that
 modulates the activity of a protease, by contacting the protease with a
 test compound, measuring the activity of the protease before and after
 contacting step and determining whether test compound modulates the
 activity of the protease. The invention also relates to the proteins not
 previously identified as proteases. The method is useful for identifying
 a compound that modulates the activity of a protease. It is useful for
 treating a disease or disorder e.g., cancer, immune-related diseases and
 disorders, cardiovascular diseases, brain or neuronal-associated diseases
 or metabolic disorders (preferably cancer of tissue or haematopoietic
 origin) in mammal, preferably human. The method is also useful for
 treating diseases of the peripheral nervous system, Alzheimer's disease,
 Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 viral infections caused by prions, infections caused by bacteria and
 fungi, ocular diseases, migraines, pain, sexual dysfunction, mood
 disorders, attention disorders, cognition disorders, hypotension,
 hypertension, psychotic disorders, neurological disorders, dyskinesias
 and organ transplant rejection. The present sequence is the human
 cysteine peptidase, a novel protease enzyme.
 Sequence 522 AA;
 Query Match 100.0%; Score 412; DB 8; Length 522;
 Best Local Similarity 100.0%; Pred. No. 2.8e-39;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SDVYCEVCEFLVKVETKLDNNKTEKILDAFKMCKLPSLSSECQVVDYTGSSILS 60
 DB 309 SDVYCEVCEFLVKVETKLDNNKTEKILDAFKMCKLPSLSSECQVVDYTGSSILS 368
 QY 61 ILLEEVSPVLVCSMLHLCSG 80
 DB 369 ILLEEVSPVLVCSMLHLCSG 388
 RESULT 13
 AAB31916
 ID AAB31916 standard; protein; 523 AA.
 XX XX
 AC AAB31916;
 XX XX
 DT 15-MAY-2001 (first entry)
 XX XX
 DE Amino acid sequence of a human protein.
 XX XX
 KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
 KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
 KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.
 XX OS
 OS Homo sapiens.
 XX XX
 PN WO200105422-A2.
 XX XX
 PD 25-JAN-2001.
 XX XX
 PF 17-JUL-2000; 2000WO-FR002057.

us-10-801-517-2.rag

Wed Jan 18 11:58:41 2006

```

XX PR 15-JUL-1999; 99PR-00009372.
XX XX (INMR ) BIOMERIEUX STELHVS.
XX PA Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
XX PI WPI; 2001-159475/16.
XX XX Detecting, preventing and treating degenerative, neurological and
XX DR auto-immune diseases, particularly multiple sclerosis, using specified
XX PT polypeptides or related nucleic acid or ligand.
XX PS Claim 1; Page 174-175; 209pp; French.
XX XX The present sequence represents a human protein, which is used in the
XX CC method of the invention. The specification describes a method which uses
XX CC at least one polypeptide or polynucleotide sequence belonging to the
XX CC perlecan, precursor of the retinol-binding plasma protein, precursor of
XX CC the ganglioside GM2 activator, calgranulin B or saposin B protein
XX CC families. The method is used for detecting, preventing or treating a
XX CC degenerative, neurological and/or auto-immune disease. The
XX CC polynucleotides and polypeptides are used for diagnosis, prognosis,
XX CC prevention and treatment of multiple sclerosis (in its various forms and
XX CC phases). They may also be useful in cases of e.g. Alzheimer's and
XX CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
XX CC polyarthritis and lupus erythematosus, including use as vaccines and in
XX CC gene therapy (expression of sense or antisense sequences). They can also
XX CC be used to assess efficacy of potential therapeutic agents, particularly
XX CC compounds that reduce or inhibit toxicity towards glial cells
XX XX Sequence 523 AA;
XX SQ Query Match 100.0%; Score 412; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.8e-39;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 310 SDVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 369
QY 61 ILLEEVSPELVCSMLHLCSG 80
Db 370 ILLEEVSPELVCSMLHLCSG 389
RESULT 14
ABU05211
ID ABU05211 standard; protein; 523 AA.
XX AC ABU05211;
XX XX 29-JAN-2003 (first entry)
XX DT Human expressed protein tag (EPT) #1877.
XX DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX KW protease; protease inhibitor; transporter; cytoskeletal protein;
XX KW receptor; transcription factor; cancer; MHC;
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX XX WO200278524-A2.
XX PN 10-OCT-2002.
XX PD 28-MAR-2002; 2002WO-US009671.
XX PF 21-MAY-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.

```

```

PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX XX (ZYCO-) ZYCOS INC.
XX PA Chiciz RM, Tomlinson AJ, Urban RG;
XX PI WPI; 2003-040607/03.
XX DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX XX cytoskeletal proteins, receptors or transcription factors), useful for
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX PT leukemia.
XX XX Example 2; SEQ ID NO 1877; 134pp; English.
XX XX The invention describes a purified polypeptide, which comprises a
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The
XX CC polypeptide is useful as an immunogenic composition for eliciting in a
XX CC mammal an immunogenic response directed against any of the purified
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this
XX CC polypeptide, is useful for treating cancer. The polypeptide is also
XX CC useful for identifying compounds that binds to a naturally processed
XX CC class I or class II MHC-binding polypeptide. The polypeptides and
XX CC polynucleotides are particularly useful for treating or preventing
XX CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX CC lymphoma or leukaemia. These are also useful for screening agents for
XX CC treating the above mentioned diseases. This sequence represents an
XX CC expressed protein tag (EPT) isolated from human tissue for translational
XX CC profiling. Note: This sequence does not appear in the printed from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX Sequence 523 AA;
XX SQ Query Match 100.0%; Score 412; DB 6; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.8e-39;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 310 SDVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 369
QY 61 ILLEEVSPELVCSMLHLCSG 80
Db 370 ILLEEVSPELVCSMLHLCSG 389
RESULT 15
AAR70783
ID AAR70783 standard; protein; 524 AA.
XX AC AAR70783;
XX XX 25-MAR-2003 (revised)
XX DT 30-AUG-1995 (first entry)
XX DE Prosaposin.
XX KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;
XX KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;
XX KW adrenal leukodystrophy; prosaposin.
XX OS Homo sapiens.
XX XX WO9503821-A1.
XX PN 09-FEB-1995.
XX PD 28-JUL-1994; 94WO-US008453.
XX PF 28-JUL-1994; 94WO-US008453.
XX XX

```

PR 30-JUL-1993; 93US-00100247.
XX 21-APR-1994; 94US-00232513.
XX (OBRI/) O'BRIEN J S.
XX O'brien JS, Kishimoto Y;
XX PI
XX DR WPI; 1995-082029/11.
XX DR N-PSDB; AAQ85355.
XX
PT Stimulating neural cell out-growth and myelination - with pro.saposin,
PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating
PT nervous system diseases.
XX
XX Disclosure; Page 30-32; 50pp; English.
XX
XX The peptide given in AAR70773, corresponding to amino acids 8-29 of human
XX saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus
XX sequence was determined by comparing the peptide with hematopoietic and
XX neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were
XX identified in the AB loop of human ciliary neurotrophic factor,
XX interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte
XX inhibitory factor, and in helix C of human interleukin-1-beta and
XX oncostatin-M. Prosapoinin (AAR70783) and saposin-C also promoted nerve
XX cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
XX correct PI field.)
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 412; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. NO. 2.8e-39;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370

Qy 61 ILLEVSPELVCSMLHLCSG 80
Db 371 ILLEVSPELVCSMLHLCSG 390

Search completed: January 13, 2006, 16:30:19
Job time : 27.947 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:22:54 ; Search time 30.4636 Seconds
(without alignments)
1852.779 Million cell updates/sec

Title: US-10-801-517-2

Perfect score: 412

Sequence: 1 SDVYCEVCEFLVKEVTKLID.....ILLREVSPELVCSMLHLCSG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412	100.0	452	2	QSR406 pongo pygma
2	412	100.0	524	1	SAP_HUMAN
3	412	100.0	524	2	Q53FJ5 homo sapien
4	412	100.0	524	2	Q5JQ36 HUMAN
5	412	100.0	527	2	Q5NVD5_PONPY
6	412	100.0	530	2	Q59EN5_HUMAN
7	412	100.0	559	2	Q5JQ37_PONPY
8	405	98.3	526	2	QSR4U7_PONPY
9	397	96.4	497	2	Q4R590_MACFA
10	346	84.0	525	1	SAP_BOVIN
11	261	63.3	81	1	SAP_CAVPO
12	210	51.0	512	2	Q7SY70_XENLA
13	207	50.2	550	2	Q4RQ38_TETNG
14	205	49.8	520	2	Q8UV24_BRARE
15	205	49.8	520	2	Q6P3G7_BRARE
16	205	49.8	520	2	Q6P3G7_BRARE
17	205	49.8	522	2	Q6P3G7_BRARE
18	203	49.3	518	1	SAP_CHICK
19	193	46.8	518	2	Q64286_XENLA
20	193	46.8	553	2	Q6P7A4_RAT
21	193	46.8	554	1	SAP_RAT
22	171	41.5	200	2	Q75K05_DICDI
23	169	41.0	554	2	Q8BFQ1_MOUSE
24	169	41.0	557	1	SAP_MOUSE
25	164	39.8	522	2	Q54Q68_DICDI
26	163	39.6	336	2	Q54LG3_DICDI
27	158	38.3	965	2	Q15937_BOMMO
28	155	37.6	245	2	Q5ZL62_CHICK
29	153	37.1	238	2	Q5BZ62_SCHJA
30	153	37.1	922	2	Q5D981_SCHJA
31	150	36.4	876	2	Q8IMH4_DROME

32 150 36.4 953 2 Q9Y125 DROME
33 149.5 36.3 253 2 QSR848_PONPY
34 149 36.2 890 2 Q7PMW6_ANOGA
35 147.5 35.8 241 2 Q8N7T4_HUMAN
36 147.5 35.8 325 2 Q7PJB6_ANOGA
37 147.5 35.8 478 2 Q7PJB7_ANOGA
38 147.5 35.8 543 2 Q6NUJ1_HUMAN
39 144 35.0 262 2 Q4SJ83_TETNG
40 144 35.0 441 2 Q5U9P3_DROME
41 143 34.7 240 2 Q5BJH1_HUMAN
42 142 34.5 293 2 Q54PT7_DICDI
43 141.5 34.3 402 2 Q8C1N0_MOUSE
44 141.5 34.3 449 2 Q8BJV5_MOUSE
45 141.5 34.3 525 2 Q8C1C1_MOUSE

Q9Y125 drosophila
Q5E848 pongo pygma
Q7PMW6 anopheles g
Q8N7T4 homo sapien
Q7PJB6 anopheles g
Q7PJB7 anopheles g
Q6NUJ1 homo sapien
Q4SJ83 tetraodon n
Q5U9P3 drosophila
Q5BJH1 homo sapien
Q54PT7 dictyosteli
Q8C1N0 mus musculu
Q8BJV5 mus musculu
Q8C1C1 mus musculu

ALIGNMENTS

RESULT 1
QSR406_PONPY PRELIMINARY; PRT; 452 AA.
AC QSR406;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459G152.
GN Name=DKFZp459G152;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobio G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR861454; CAH93510.1; -, mRNA.
DR SMR; QSR406; 119-200, 239-318.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa.1.
DR InterPro; IPR008138; Sapa.2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; Saposinb.
DR Pfam; PF02199; Sapa.1.
DR Pfam; PF05184; Sapa.1; 4.
DR Pfam; PF03489; Sapa.2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; Sapa_sub.2.
DR SMART; SM00162; SAPA; 1.
DR KW Hypothetical protein.
SQ SEQUENCE 452 AA; 50664 MW; 6D30EAC665A399DD CRC64;

Query Match 100.0%; Score 412; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.2e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 60
|||||
|||||

Db 239 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 298
|||||
|||||

Qy 61 ILLEVSPELVCSMLHLCSG 80
|||||
|||||

Db 299 ILLEVSPELVCSMLHLCSG 318
|||||
|||||

Wed Jan 18 11:58:43 2006

RESULT 2

SAP_HUMAN STANDARD; PRT; 524 AA.

ID SAL_HUMAN P07602; P15793; P78538; P78541; P78546; P78547; P78558;

AC Q61BQ6; Q92739; Q92740; Q92741; Q92742;

AC Q61BQ6; Q92739; Q92740; Q92741; Q92742;

DT 01-APR-1998 (Rel. 07, Created)

DT 01-APR-1998 (Rel. 14, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);

DE Saposin B-val; Saposin B (Sphingolipid activator protein 1) (SAP-1)

DE (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1

DE activator); Saposin C (Co-beta-glucosidase) (AI activator)

DE (Glucosylceramide activator) (Sphingolipid activator protein 2)

DE (SAP-2); Saposin D (Protein C) (Component C)].

GN Name=PSAP;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

[1]

RN NUCLEOTIDE SEQUENCE [MRNA].

RP TISSUE=Liver;

RC MEDLINE=90129043; PubMed=2515150;

RX Rorman E.G., Grabowski G.A.;

RA "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that

RT four sphingolipid hydrolase activator proteins are encoded by single

RT genes in humans and rats.;"

RL J. Biochem. 105:152-154 (1989).

[2]

RN NUCLEOTIDE SEQUENCE [MRNA].

RP MEDLINE=8925151; PubMed=2498298;

RA Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;

RA "Structure of full-length cDNA coding for sulfatide activator, a Co-

RT beta-glucosidase and two other homologous proteins: two alternate

RT forms of the sulfatide activator.;"

RL J. Biochem. 105:152-154 (1989).

[3]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RP Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;

RA "Cloning of human full open reading frames in Gateway(TM) system entry

RT vector (pDONR201).;"

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

[4]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SAP-MU-0).

RP TISSUE=Brain, Eye, and Skin;

RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zerbey B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzyzaniak M.I., Skalska U.,

RA Sanger A., Schein J.E., Jones S.J.M., Marra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.;"

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[5]

RN NUCLEOTIDE SEQUENCE OF 14-524.

RP MEDLINE=88321660; PubMed=2842863;

RA O'Brien J.S., Kretz K.A., Dewji N., Wenger D.A., Esch F.,

RA Fluhrty A.L.;

RT "Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by

RT same genetic locus.;"

RL Science 241:1098-1101 (1988).

[6]

RN NUCLEOTIDE SEQUENCE OF 14-524.

RP MEDLINE=92307663; PubMed=1612590;

RA Rorman E.G., Scheinker V., Grabowski G.A.;

RA "Structure and evolution of the human prosaposin chromosomal gene.;"

RL Genomics 13:312-318 (1992).

[7]

RN PROTEIN SEQUENCE OF 17-24; 165-172; 180-189 AND 298-302.

RP MEDLINE=93311991; PubMed=8323276; DOI=10.1006/abbi.1993.1328;

RA Hiraiwa M., O'Brien J.S., Kishimoto Y., Galdzicka M., Fluhrty A.L.,

RA Gims E.I., Martin B.M.;

RA "Isolation, characterization, and proteolysis of human prosaposin, the

RT precursor of saposins (sphingolipid activator proteins).;"

RL Arch. Biochem. Biophys. 304:110-116 (1993).

[8]

RN PROTEIN SEQUENCE OF 17-26.

RP TISSUE=Milk;

RC MEDLINE=92068206; PubMed=1958198;

RA Kondoh K., Hinenio T., Sano A., Kakimoto Y.;

RA "Isolation and characterization of prosaposin from human milk.;"

RL Biochem. Biophys. Res. Commun. 181:286-292 (1991).

[9]

RN NUCLEOTIDE SEQUENCE OF 59-125 AND 304-513.

RP TISSUE=Brain;

RC MEDLINE=93192146; PubMed=2013321; DOI=10.1016/0014-5793(91)80308-P;

RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,

RA Suzuki K.;

RA "The organization of the gene for the human cerebroside sulfate

RT activator protein.;"

RL FEBS Lett. 280:267-270 (1991).

[10]

RN PARTIAL PROTEIN SEQUENCE OF 60-142.

RP MEDLINE=89240739; PubMed=2717620;

RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,

RA Kishimoto Y.;

RA "Saposin A: second cerebroside activator protein.;"

RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393 (1989).

[11]

RN PROTEIN SEQUENCE OF 62-84 AND 410-431.

RP MEDLINE=93380576; PubMed=8370464; DOI=10.1016/0014-5793(93)80908-D;

RA Tyne J., Palmer D.N., Baumann M., Haila M.;

RA "Storage of saposins A and D in infantile neuronal ceroid-

RT lipofuscinosis.;"

RL FEBS Lett. 330:8-12 (1993).

[12]

RN NUCLEOTIDE SEQUENCE OF 164-524.

RP MEDLINE=88068647; PubMed=2825202;

RA Dewji N.N., Wenger D.A., O'Brien J.S.;

RA "Nucleotide sequence of cloned cDNA for human sphingolipid activator

RT protein 1 precursor.;"

RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656 (1987).

[13]

RN NUCLEOTIDE SEQUENCE OF 195-263.

RP MEDLINE=86130593; PubMed=2868718;

RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,

RA Hill F., O'Brien J.S.;

RA "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),

RT the sulfatide sulfate activator.;"

RL Biochem. Biophys. Res. Commun. 134:989-994 (1986).

[14]

RN PROTEIN SEQUENCE OF 195-274.

RP MEDLINE=89207118; PubMed=23242555;

RA Kleinschmidt T., Christomanou H., Braunitzer G.;

RA "Complete amino-acid sequence of the naturally occurring A2 activator

RT protein for enzymic sphingomyelin degradation: identity to the

RT sulfatide activator protein (SAP-1).;"

RL Biol. Chem. Hoppe-Seyler 369:1361-1365 (1988).

[15]

RN PROTEIN SEQUENCE OF 195-274.

RP TISSUE=Kidney;

RC MEDLINE=91006165; PubMed=2209618;

RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
RT "The complete amino-acid sequences of human ganglioside GM2 activator
RL protein and cerebroside sulfate activator protein.";
RL Eur. J. Biochem. 192:709-714(1990).
RN [16]
RN PROTEIN SEQUENCE OF 311-390.
RX MEDLINE=88163077; PubMed=3442600;
RA Kleinschmidt T., Christomanou H., Braunitzer G.;
RT "Complete amino-acid sequence and carbohydrate content of the
RT naturally occurring glucosylceramide activator protein (A1 activator)
RT absent from a new human Gaucher disease variant.";
RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
RN [17]
RN PARTIAL PROTEIN SEQUENCE OF 405-484.
RX MEDLINE=89025876; PubMed=2845979;
RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
RT "Saposin D: a sphingomyelinase activator.";
RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
RN [18]
RN PROTEIN SEQUENCE OF 407-484.
RX MEDLINE=89000190; PubMed=3048308;
RA Furst W., Machleidt W., Sandhoff K.;
RT "The precursor of sulfatide activator protein is processed to three
RT different proteins.";
RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
RN [19]
RN PARTIAL PROTEIN SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
RC TISSUE=Urine;
RX MEDLINE=20032116; PubMed=10562467; DOI=10.1006/mgme.1999.2900;
RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,
RA Waring A.J., To T., Fluharty C.B., Faull K.F.;
RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)
RT from human urine.";
RL Mol. Genet. Metab. 68:391-403(1999).
RN [20]
RN DISULFIDE BONDS OF SAPOSINS B AND C, AND MASS SPECTROMETRY.
RX MEDLINE=95247790; PubMed=730378; DOI=10.1074/jbc.270.17.9953;
RA Vaccaro A.M., Salviole R., Barca A., Tatti M., Claffoni F., Maras B.,
RA Siciliano R., Zappacosta P., Amoresano A., Pucci P.;
RT "Structural analysis of saposin C and B. Complete localization of
RT disulfide bridges.";
RL J. Biol. Chem. 270:9953-9960(1995).
RN [21]
RN STRUCTURE OF CARBOHYDRATE ON ASN-215.
RX MEDLINE=21110404; PubMed=11180632;
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN [22]
RN SAPOSIN D DISULFIDE BONDS.
RX MEDLINE=99337688; PubMed=10406958;
RA Tatti M., Salviole R., Claffoni F., Pucci P., Andolfo A.,
RA Amoresano A., Vaccaro A.M.;
RT "Structural and membrane-binding properties of saposin D.";
Query Match 100.0%; Score 412; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. NO. 2.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCKSLPKSLSEECQEVVDYTGSSILS 60
DB 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCKSLPKSLSEECQEVVDYTGSSILS 370
QY 61 ILLEEVSPVLCSMLHLCSG 80
DB 371 ILLEEVSPVLCSMLHLCSG 390
RESULT 3
Q53FJ5 HUMAN
ID Q53FJ5_HUMAN PRELIMINARY; PRT; 524 AA.

AC Q53FJ5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Prosaposin (Variant Gaucher disease and variant metachromatic
DE leukodystrophy) variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Suzuki Y., Yoshitomo K., Maruyama K., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [3]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Tanaka A., Yokoyama S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK232290; BAD97010.1; -; mRNA.
FT NON_TER 1
SQ SEQUENCE 524 AA; 58140 MW; 7034F0C71C2226BC CRC64;
Query Match 100.0%; Score 412; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. NO. 2.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCKSLPKSLSEECQEVVDYTGSSILS 60
DB 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCKSLPKSLSEECQEVVDYTGSSILS 370
QY 61 ILLEEVSPVLCSMLHLCSG 80
DB 371 ILLEEVSPVLCSMLHLCSG 390
RESULT 4
Q5JQ36 HUMAN
ID Q5JQ36_HUMAN PRELIMINARY; PRT; 524 AA.
AC Q5JQ36;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Prosaposin (Variant Gaucher disease and variant metachromatic
DE leukodystrophy).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor

Wed Jan 18 11:58:43 2006

```
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731541; CAI40837.1; -; Genomic_DNA.
DR EMBL; BT006849; AAP35495.1; -; mRNA.
DR EMBL; AC073370; CAI40837.1; JOINED; Genomic_DNA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
SQ SEQUENCE 524 AA; 58112 MW; 71977F7A8C9B1533 CRC64;

Query Match 100.0%; Score 412; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 60
Db 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 370

Qy 61 ILLEEVSPVLCSMLHLCSG 80
Db 371 ILLEEVSPVLCSMLHLCSG 390

RESULT 5
Q5NVDS5_PONPY PRELIMINARY; PRT; 527 AA.
AC Q5NVDS5_PONPY PRELIMINARY; PRT; 527 AA.
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459F0110.
GN Name=DKFZp459F0110;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Cortex;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CE926102; CAI29728.1; -; mRNA.
DR SMR; Q5NVDS5; 194-275, 314-393.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SApB.1.
DR InterPro; IPR008138; SApB.2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SApB_1; 4.
DR Pfam; PF03489; SApB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SApB_sub; 2.
DR SMART; SM00162; SAPA; 2.
KW Hypothetical protein.
SQ SEQUENCE 527 AA; 58469 MW; 293FBB746C29C4D0 CRC64;

Query Match 100.0%; Score 412; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 60
Db 314 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 373

Qy 61 ILLEEVSPVLCSMLHLCSG 80
```

```
Db 374 ILLEEVSPVLCSMLHLCSG 393

RESULT 6
Q59EN5_HUMAN PRELIMINARY; PRT; 530 AA.
AC Q59EN5_HUMAN PRELIMINARY; PRT; 530 AA.
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Prosaposin variant (Fragment).
GN Name=prosaposin variant;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209776; BAD93013.1; -; mRNA.
FT NON TER 1
SQ SEQUENCE 530 AA; 58727 MW; 6CA1F0159B182BC9 CRC64;

Query Match 100.0%; Score 412; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 60
Db 317 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECEQVVDVTYGSILS 376

Qy 61 ILLEEVSPVLCSMLHLCSG 80
Db 377 ILLEEVSPVLCSMLHLCSG 396

RESULT 7
Q5JQ37_HUMAN PRELIMINARY; PRT; 559 AA.
AC Q5JQ37_HUMAN PRELIMINARY; PRT; 559 AA.
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Prosaposin (Variant Gaucher disease and variant metachromatic
DE leukodystrophy).
GN Name=PSAP; ORFNames=RP11-472K8.1-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731541; CAI40836.1; -; Genomic_DNA.
DR EMBL; AC073370; CAI40836.1; JOINED; Genomic_DNA.
DR SMR; Q5JQ37; 194-272, 311-390.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SApB.1.
DR InterPro; IPR008138; SApB.2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008139; Saposin.
DR InterPro; IPR008139; SaposinB.
```

DR Pfam; PF02199; SapA; 2.
 DR Pfam; PF05184; SapB_1; 4.
 DR Pfam; PF03489; SapB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 2.
 DR SMART; SM00162; SApA; 2.
 SQ SEQUENCE 559 AA; 61692 MW; 7C80741B6039AE38 CRC64;

Query Match 100.0%; Score 412; DB 2; Length 559;
 Best Local Similarity 100.0%; Pred. No. 2.7e-30;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 60
 Db 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 370
 Qy 61 ILLEEVSPVLCSMLHLCSSG 80
 Db 371 ILLEEVSPVLCSMLHLCSSG 390

RESULT 8
 Q5RAU7_PONPY
 ID Q5RAU7_PONPY PRELIMINARY; PRT; 526 AA.
 AC Q5RAU7;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein DKFp459J1314.
 GN Names=DKFp459J1314;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Pongo.
 NCBI_TaxID=9600;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Cortex;
 RG The German cDNA Consortium;
 RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
 RA Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR861144; CAH93219.1; -; mRNA.
 DR SMR; Q5RAU7; 194-274, 313-392.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006629; P:lipid metabolism; IEA.
 DR GO; GO:0006865; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR003119; SapA.
 DR InterPro; IPR007856; SapB_1.
 DR InterPro; IPR008138; SapB_2.
 DR InterPro; IPR008140; SapB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF02199; SapA; 2.
 DR Pfam; PF05184; SapB_1; 4.
 DR Pfam; PF03489; SapB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 2.
 DR SMART; SM00162; SApA; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 526 AA; 58325 MW; CF3B146DDB6F539 CRC64;

Query Match 98.3%; Score 405; DB 2; Length 526;
 Best Local Similarity 98.8%; Pred. No. 1.2e-29;
 Matches 79; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 60
 Db 313 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 372
 Qy 61 ILLEEVSPVLCSMLHLCSSG 80
 Db 373 ILLEEVSPVLCSMLHLCSSG 392

RESULT 9

Q4R590_MACFA
 ID Q4R590_MACFA PRELIMINARY; PRT; 497 AA.
 AC Q4R590;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Brain cDNA, clone: QccE-13090, similar to human prosaposin (variant
 DE Gaucher disease and variantmetachromatic leukodystrophy) (PSAP).
 DE (Brain cDNA, clone: QccE-13989, similar to human prosaposin (variant
 DE Gaucher disease and variantmetachromatic leukodystrophy)
 DE (PSAP).
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecidae; Cercopithecinae; Macaca.
 NCBI_TaxID=9541;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA International consortium for macaque cDNA sequencing, analysis;
 RT "DNA sequences of macaque genes expressed in brain or testis and its
 RT evolutionary implications.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
 RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
 RT "Substitution rate and structural divergence of 5'UTR evolution:
 RT Comparative analysis between human and cynomolgus monkey cDNAs.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB169654; BAE01735.1; -; mRNA.
 DR EMBL; AB169527; BAE01609.1; -; mRNA.
 SQ SEQUENCE 497 AA; 55306 MW; 4A1974F8DB883900 CRC64;

Query Match 96.4%; Score 397; DB 2; Length 497;
 Best Local Similarity 96.2%; Pred. No. 6.3e-29;
 Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 60
 Db 314 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSEKLPKSLSEECQEVVDYVYSSILS 373
 Qy 61 ILLEEVSPVLCSMLHLCSSG 80
 Db 374 ILLEEVSPVLCSMLHLCSSG 393

RESULT 10
 SAP_BOVIN
 ID SAP_BOVIN STANDARD; PRT; 525 AA.
 AC R26779; Q9NZG4;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
 DE sulfate activator) (CSact)(dispersin) (Sulfatide/GM1 activator);
 DE Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
 DE (Protein C) (Component C)].
 GN Name=SAP;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND VARIANTS.
 RC TISSUE=Mammary gland;
 RA Azuma N., Yoshida K.;
 RT "RT-PCR cloning of bovine prosaposin.";

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

[2]
 PROTEIN SEQUENCE OF 312-391.
 TISSUE=Spleen;
 MEDLINE=92207994; PubMed=1554743; DOI=10.1016/0167-4838(92)90426-B;
 Sano A., Mizuno T., Kondoh K., Hineno T., Ueno S.-I., Kakimoto Y.,
 Morita N.;
 "Saposin-C from bovine spleen; complete amino acid sequence and
 relation between the structure and its biological activity.";
 Biochim. Biophys. Acta 1120:75-80(1992).
 CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place
 by the sequential action of specific hydrolases. Some of these
 enzymes require specific low-molecular mass, non-enzymic proteins:
 the sphingolipids activator proteins (coproteins) (By similarity).
 CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 lipid to form an activated complex, rather than by solubilizing
 the substrate.
 CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
 gangliosides by beta-galactosidase (EC 3.2.1.23) and
 globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 sphingolipid hydrolases (By similarity).
 CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
 activator (EC 3.1.4.12) (By similarity).
 CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
 CC -1- PTM: This precursor is proteolytically processed to 4 small
 peptides, which are similar to each other and are sphingolipid
 hydrolase activator proteins (By similarity).
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

 EMBL; AB036791; BAA95677.1; -; mRNA.
 DR HSPP; Q92739; IM12.
 DR SMR; P26779; 195-273, 312-390.
 DR InterPro; IPR009007; Pept_Aspart_cat.
 DR InterPro; IPR003119; Sapa.
 DR InterPro; IPR007856; SapB_1.
 DR InterPro; IPR008138; SapB_2.
 DR InterPro; IPR008140; SapB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR InterPro; IPR008137; Surfactant_B.
 DR Pfam; PF02199; Sapa_2.
 DR Pfam; PF05184; SapB_1; 4.
 DR Pfam; PF03489; SapB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 2.
 DR ProDom; PD008002; Surfactant_B; 1.
 DR SMART; SM00162; SAPA; 2.
 DR SMART; SM00741; SapB; 4.
 DR PROSITE; PS51110; SAPA; 2.
 DR PROSITE; PS50015; SAP B; 4.
 DR Direct protein sequencing; Glycoprotein; Lipid metabolism; Lysosome;
 KW Repeat; Signal; Sphingolipid metabolism.
 FT SIGNAL 1 16
 FT PROPEP 17 58 Saposin A.
 FT CHAIN 60 142
 FT PROPEP 144 195
 FT CHAIN 196 275 Saposin B.
 FT PROPEP 277 310
 FT CHAIN 312 392 Saposin C.
 FT PROPEP 393 404

FT CHAIN 406 487
 FT PROPEP 489 525
 FT DOMAIN 18 58
 FT DOMAIN 59 142
 FT DOMAIN 194 276
 FT DOMAIN 312 392
 FT DOMAIN 406 487
 FT DOMAIN 489 525
 FT CARBOHYD 80 80
 FT CARBOHYD 101 101
 FT CARBOHYD 216 216
 FT CARBOHYD 333 333
 FT CARBOHYD 427 427
 FT DISULFID 63 138
 FT DISULFID 66 132
 FT DISULFID 94 106
 FT DISULFID 199 272
 FT DISULFID 202 266
 FT DISULFID 231 242
 FT DISULFID 316 389
 FT DISULFID 319 383
 FT DISULFID 347 358
 FT DISULFID 410 483
 FT DISULFID 413 477
 FT DISULFID 441 452
 FT VARIANT 127 127
 FT VARIANT 260 263
 FT CONFLICT 317 317
 FT CONFLICT 367 367
 SQ SEQUENCE 525 AA; 58120 MW; 293AFC0F9C4FA99 CRC64;

Query Match 84.0%; Score 346; DB 1; Length 525;
 Best Local Similarity 81.0%; Pred. No. 4.1e-24;
 Matches 64; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKYVTKYKIDNNKTEKILDAFKWCKLPSLSSECCQWVDYTGSSILS 60
 312 ADIYCEVCEFLVKYVTKYKIDNNKTEKILDAFKWCKLPSLSSECCQWVDYTGSSILS 371
 QY 61 ILLEEVSPVLVCSMLHLS 79
 372 ILLDEASPELVCSMLHLS 390
 DB

RESULT 11
 SAP CAVPO STANDARD; PRT; 81 AA.
 AC P20097;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Saposin C (CO-beta-glucosidase) (Glucosylceramidase activator)
 DE (Sphingolipid activator protein 2) (SAP-2).
 GN Name=PSAP;
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 OC Hystricognathi; Caviidae; Cavia.
 OC NCBI_TaxID=10141;
 RN [1]
 RP PROTEIN SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=89066787; PubMed=3198642;
 RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
 RT "The activator protein for glucosylceramide beta-glucosidase from
 guinea pig liver. Improved isolation method and complete amino acid
 sequence.";
 RT J. Biol. Chem. 263:19597-19601(1988).
 RL -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 lipid to form an activated complex, rather than by solubilizing

```
CC the substrate.
CC 1- SIMILARITY: Contains 1 saposin B-type domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PIR: A32026; A32026.
CC HSSP: Q92739; IM12.
CC InterPro: IPR007856; SapB 1.
CC InterPro: IPR008138; SapB 2.
CC InterPro: IPR008140; SapB sub.
CC InterPro: IPR008373; Saposin.
CC InterPro: IPR008139; SaposinB.
CC Pfam: PF05184; SapB 1; 1.
CC Pfam: PF03489; SapB 2; 1.
CC PRINTS: PR01797; SAPOSIN.
CC ProDom: PD001732; SapB sub; 1.
CC SMART: SM00741; SapB 1.
CC PROSITE: PS00015; SAP_B; 1.
CC Direct protein sequencing; Glycoprotein; Lipid metabolism;
KW Sphingolipid metabolism.
CC
CC FT DOMAIN 1 81 Saposin B-type.
CC CARBOHYD 22 22 N-linked (GlcNAc . . ).
CC DISULFID 5 78 By similarity.
CC FT DISULFID 8 72 By similarity.
CC FT DISULFID 36 47 By similarity.
CC SEQUENCE 81 AA; 8852 MW; E564CE1F0A292596 CRC64;
SQ
Query Match 63.3%; Score 261; DB 1; Length 81;
Best Local Similarity 61.8%; Pred. No. 6.3e-17;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;
QY 3 VYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECEQVVDYTGSSILSIL 62
DB 3 VTCKACEYVVKWELIDNNRTREKIIHALDSVALLPESVSEVCEQVVDYTGDSIVALL 62
QY 63 LIEVSPVLCVSMHLHC 78
DB 63 LQEMSPVLCVSELGLC 78
RESULT 12
Q7SY70_XENLA PRELIMINARY; PRT; 512 AA.
AC Q7SY70;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MGC64541 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshitoki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
```

```
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054988; AAHS4988.1; -; mRNA.
DR HSSP; Q92739; IM12.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro: IPR003119; SapA.
DR InterPro: IPR007856; SapB 1.
DR InterPro: IPR008138; SapB 2.
DR InterPro: IPR008140; SapB sub.
DR InterPro: IPR008373; Saposin.
DR InterPro: IPR008139; SaposinB.
DR Pfam: PF02199; SapA; 2.
DR Pfam: PF05184; SapB 1; 4.
DR Pfam: PF03489; SapB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB sub; 2.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00741; SapB; 4.
SQ SEQUENCE 512 AA; 57555 MW; 57CPA5E2093F6PB1 CRC64;
Query Match 51.0%; Score 210; DB 2; Length 512;
Best Local Similarity 37.5%; Pred. No. 2.4e-11;
Matches 30; Conservative 27; Mismatches 23; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECEQVVDYTGSSILS 60
DB 299 NNVCYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECEQVVDYTGSSILS 358
QY 61 ILLEVSPELVCSMLHLCSSG 80
DB 359 LLEQEANPEVICTTLGYCSG 378
RESULT 13
Q4RQ38_TETNG PRELIMINARY; PRT; 550 AA.
AC Q4RQ38;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 17 SCAR15006, whole genome shotgun sequence.
DE (fragment).
GN ORFNames=GSTENG00030786001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```


DR Pfam; PF05184; SapB_1; 3.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 2.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00741; SapB; 4.
SQ SEQUENCE 520 AA; 57431 MW; F9E620F94BA41CB5 CRC64;

Query Match 49.8%; Score 205; DB 2; Length 520;
Best Local Similarity 40.8%; Pred. No. 7.3e-11;
Matches 31; Conservative

Oy 5 CEVCEFLVKEVTKLIDNNKTEKILDAFDQWCKLPKLSLSECEQVDVTCSSILSILLE 64
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 312 CAICEYMKIEINMIQDTSEAIVQAWEKVNCILPSTLTAAQCKDLIETYGQAIDLVLVQ 371

Oy 65 EVSPELVCMSMLHLCSG 80
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 372 EADPKTVCISFLALCSG 387

RESULT 15
O6P3G7 BRARE
ID Q6P3G7_BRARE PRELIMINARY; PRT; 520 AA.

AC Q6P3G7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Prosaposin.
GN Name=psap;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RF NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyminski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RC NUCLEOTIDE SEQUENCE.
RC TISSUE=whole;
RA Strausberg R.;
RL Submitted (DSC-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC063994; AAH63994.1; -, mRNA.
DR ZFIN; ZDB-GENE-020108-1; psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB.1.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:26:40 ; Search time 15.2318 Seconds
(without alignments)
2194.512 Million cell updates/sec

Title: US-10-801-517-2
Perfect score: 412
Sequence: 1 SDVYCEVFLVKEVTKLID.....ILLEVSPVLVCSMLHLCSSG 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pap:*
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pap:*
4: /cgn2_6/prodata1/pubpaa/US10_PUBCOMB.pap:*
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pap:*
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	100.0	80	3	US-09-767-007A-3
2	412	100.0	80	3	US-09-753-126-3
3	412	100.0	80	4	US-10-330-697-3
4	412	100.0	80	4	US-10-746-442-24
5	412	100.0	80	5	US-10-801-517-2
6	412	100.0	80	6	US-11-036-867-24
7	412	100.0	210	5	US-10-473-127-1867
8	412	100.0	479	3	US-09-978-418-40
9	412	100.0	479	5	US-10-485-231-40
10	412	100.0	522	5	US-10-618-281-38
11	412	100.0	523	3	US-09-767-007A-2
12	412	100.0	523	5	US-10-473-127-1877
13	412	100.0	524	3	US-09-870-759-60
14	412	100.0	524	3	US-09-751-708A-60
15	412	100.0	524	4	US-10-267-502-386
16	412	100.0	524	4	US-10-408-765A-1207
17	412	100.0	524	4	US-10-746-442-23
18	412	100.0	524	4	US-10-428-817A-56
19	412	100.0	524	5	US-10-801-517-1
20	412	100.0	524	5	US-10-473-127-1865
21	412	100.0	524	5	US-10-473-127-1866
22	412	100.0	524	5	US-10-473-127-1868
23	412	100.0	524	5	US-10-473-127-1869
24	412	100.0	524	5	US-10-473-127-1871
25	412	100.0	524	5	US-10-473-127-1873
26	412	100.0	524	5	US-10-473-127-1874
27	412	100.0	524	5	US-10-473-127-1878

28	412	100.0	524	5	US-10-473-127-1879	Sequence 1879, Ap
29	412	100.0	524	5	US-10-473-127-1880	Sequence 1880, Ap
30	412	100.0	524	5	US-10-473-127-1881	Sequence 1881, Ap
31	412	100.0	524	5	US-10-473-127-1882	Sequence 1882, Ap
32	412	100.0	524	5	US-10-473-127-2041	Sequence 2041, Ap
33	412	100.0	524	6	US-11-036-867-23	Sequence 23, Appl
34	412	100.0	526	3	US-09-870-759-61	Sequence 1875, Ap
35	412	100.0	527	3	US-09-751-708A-61	Sequence 61, Appl
36	412	100.0	527	3	US-10-060-036-73	Sequence 73, Appl
37	412	100.0	527	4	US-10-428-817A-57	Sequence 57, Appl
38	412	100.0	527	5	US-10-473-127-1870	Sequence 1870, Ap
39	412	100.0	527	5	US-10-473-127-1872	Sequence 1872, Ap
40	412	100.0	527	5	US-10-473-127-1876	Sequence 1876, Ap
41	412	100.0	592	3	US-09-753-126-4	Sequence 4, Appl
42	412	100.0	592	4	US-10-330-697-4	Sequence 4, Appl
43	412	100.0	592	4	US-10-452-858C-11	Sequence 11, Appl
44	355	86.2	227	4	US-10-043-487-340	Sequence 340, App
45	257	62.4	209	4		

ALIGNMENTS

RESULT 1

US-09-767-007A-3
; Sequence 3, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; TITLE OF INVENTION: SAPOSIN C AND NEUTROTROPIC PEPTIDES DERIVED THEREFROM
; FILE REFERENCE: MYELOS 2DC1C1
; CURRENT APPLICATION NUMBER: US/09767,007A
; PRIORITY FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide
US-09-767-007A-3

Query Match	100.0%	Score 412;	DB 3;	Length 80;
Best Local Similarity	100.0%	Pred. No. 8.4e-37;		
Matches	80;	Conservative	0;	Mismatches
		Indels	0;	Gaps
				0;
Qy	1	SDVYCEVFLVKEVTKLIDNNKTEKILDA	PKSKLPKSLSECCQEVVDYSSILS	60
Db	1	SDVYCEVFLVKEVTKLIDNNKTEKILDA	PKSKLPKSLSECCQEVVDYSSILS	60
Qy	61	ILLEEVSPVLVCSMLHLCSSG	80	
Db	61	ILLEEVSPVLVCSMLHLCSSG	80	

RESULT 2

US-09-753-126-3
; Sequence 3, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-000600US
; CURRENT APPLICATION NUMBER: US/09753,126

Wed Jan 18 11:58:42 2006

```

; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/217,497
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-753-126-3

```

```

Query Match      100.0%; Score 412; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-37;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60
    |||||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60

QY 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 61 ILLEEVSPELVCSMLHLCSG 80

```

```

RESULT 3
US-10-330-697-3
; Sequence 3, Application US/10330697
; Publication No. US2004009165A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-000600US
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-697-3

```

```

Query Match      100.0%; Score 412; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-37;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60
    |||||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60

QY 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 61 ILLEEVSPELVCSMLHLCSG 80

```

```

RESULT 4
US-10-746-442-24
; Sequence 24, Application US/10746442
; Publication No. US20040121958A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; FILE REFERENCE: 07256/024001
; CURRENT APPLICATION NUMBER: US/10/746,442
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US/08/928,074
; PRIOR FILING DATE: FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 08/611,307
; PRIOR FILING DATE: FILING DATE: 1996-03-05
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: PCT/US97/04143
; PRIOR FILING DATE: FILING DATE: 1996-03-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
; OTHER INFORMATION: Peptide Sequence
US-10-746-442-24

```

```

Query Match      100.0%; Score 412; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-37;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60
    |||||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMKSLPKSLSECCQEVVDYTGSSILS 60

QY 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 61 ILLEEVSPELVCSMLHLCSG 80

```

```

RESULT 5
US-10-801-517-2
; Sequence 2, Application US/10801517
; Publication No. US20040229799A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Xiaoyang
; APPLICANT: Children's Hospital Medical Center
; TITLE OF INVENTION: Saposin C-DOPS A Novel Anti-Tumor Agent
; FILE REFERENCE: CHM08/GN003
; CURRENT APPLICATION NUMBER: US/10/801,517
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 60/466,166
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-801-517-2

Query Match 100.0%; Score 412; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-37; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 60
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 60
QY 61 ILLEEVSPVLVCSMLHLCSSG 80
DB 61 ILLEEVSPVLVCSMLHLCSSG 80

RESULT 6

US-11-036-867-24
; Sequence 24, Application US/11036867
; Publication No. US20050164948A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/036,867
; FILING DATE: 14-Jan-2005
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,074
; FILING DATE: 11-SEP-1997
; APPLICATION NUMBER: 08/611,307
; FILING DATE: 05-MAR-1996
; APPLICATION NUMBER: PCT/US97/04143
; FILING DATE: 05-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07256/024001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-11-036-867-24

Query Match 100.0%; Score 412; DB 6; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-37; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 60
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 60
QY 61 ILLEEVSPVLVCSMLHLCSSG 80
DB 61 ILLEEVSPVLVCSMLHLCSSG 80

RESULT 7

US-10-473-127-1867
; Sequence 1867, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1867
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1867

Query Match 100.0%; Score 412; DB 5; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.8e-36; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 60
DB 8 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCSECEVVDYTGSSILS 67
QY 61 ILLEEVSPVLVCSMLHLCSSG 80
DB 68 ILLEEVSPVLVCSMLHLCSSG 87

RESULT 8

US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephan
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142.USS.REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

Query Match 100.0%; Score 412; DB 3; Length 479;
Best Local Similarity 100.0%; Pred. No. 7.8e-36;

Wed Jan 18 11:58:42 2006

Query Match 100.0%; Score 412; DB 5; Length 522;
Best Local Similarity 100.0%; Pred. No. 8.6e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 60
Db 314 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 373
QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 374 ILLEEVSPELVCSMLHLCSSG 393

RESULT 9
US-10-485-231-40
; Sequence 40, Application US/10485231
; Publication No. US2005011917A1
; GENERAL INFORMATION:
; APPLICANT: GENSET
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142.WO1
; CURRENT APPLICATION NUMBER: US/10/485.231
; CURRENT FILING DATE: 2004-01-29
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Jpatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-231-40

Query Match 100.0%; Score 412; DB 5; Length 479;
Best Local Similarity 100.0%; Pred. No. 7.8e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 60
Db 314 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 373
QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 374 ILLEEVSPELVCSMLHLCSSG 393

RESULT 10
US-10-618-281-38
; Sequence 38, Application US/10618281
; Publication No. US20040215609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteases Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-38

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 60
Db 314 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 373
QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 374 ILLEEVSPELVCSMLHLCSSG 393

RESULT 11
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US2002007725A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: MYELOS.2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767.007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match 100.0%; Score 412; DB 3; Length 523;
Best Local Similarity 100.0%; Pred. No. 8.7e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 60
Db 310 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDYTGSSILS 369
QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 370 ILLEEVSPELVCSMLHLCSSG 389

RESULT 12
US-10-473-127-1877
; Sequence 1877, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 127
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1877

```
; SEQ ID NO 1877
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1877

Query Match      100.0%; Score 412; DB 5; Length 523;
Best Local Similarity 100.0%; Pred. No. 8.7e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 310 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 369

QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 370 ILLEEVSPELVCSMLHLCSSG 389

RESULT 13
US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-60

Query Match      100.0%; Score 412; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 8.7e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 311 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370

QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 371 ILLEEVSPELVCSMLHLCSSG 390

RESULT 14
US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match      100.0%; Score 412; DB 3; Length 524;
```

```
Best Local Similarity 100.0%; Pred. No. 8.7e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 311 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370

QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 371 ILLEEVSPELVCSMLHLCSSG 390

RESULT 15
US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match      100.0%; Score 412; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 8.7e-36;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
Db 311 SDVYCEVCFVLKVEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370

QY 61 ILLEEVSPELVCSMLHLCSSG 80
Db 371 ILLEEVSPELVCSMLHLCSSG 390

Search completed: January 13, 2006, 16:38:41
Job time : 15.2318 secs
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:30:30 ; Search time 1.5894 Seconds
(without alignments)
475.862 Million cell updates/sec

Title: US-10-801-517-2

Perfect score: 412
Sequence: 1 SDVCEVCEFLVKEVTKLID.....ILLEVSPELVCSMLHLCSSG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 67062 seqs, 9454214 residues

Total number of hits satisfying chosen parameters: 67062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412	100.0	524	6	US-10-821-234-1631
2	70.5	17.1	129	7	US-11-078-469-20
3	70.5	17.1	172	7	US-11-078-469-22
4	70	17.0	1786	7	US-11-196-400-3
5	67.5	16.4	145	7	US-11-078-469-21
6	65	15.8	989	6	US-10-821-234-975
7	64	15.5	476	6	US-10-878-556A-158
8	61.5	14.9	115	6	US-10-467-657-5296
9	61	14.8	292	7	US-11-194-246-316
10	60	14.6	2036	7	US-11-124-368A-276
11	60	14.6	2036	7	US-11-124-368A-280
12	60	14.6	2036	7	US-11-124-368A-281
13	59.5	14.4	3002	6	US-10-821-234-916
14	59	14.3	197	5	US-09-940-308-3
15	59	14.3	389	5	US-09-940-308-6
16	59	14.3	590	5	US-09-940-308-2
17	57.5	14.0	1738	7	US-11-078-469-19
18	57.5	14.0	1138	6	US-10-509-422-4
19	57.5	14.0	5935	6	US-10-995-561-776
20	57	13.8	1087	7	US-11-102-978-2
21	57	13.8	1400	6	US-10-821-234-1045
22	56.5	13.7	122	6	US-10-793-626-564
23	56.5	13.7	248	6	US-10-793-626-464
24	55.5	13.5	372	6	US-10-844-035-1
25	55.5	13.5	373	6	US-10-995-561-948

ALIGNMENTS

RESULT 1

US-10-821-234-1631
; Sequence 1631, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1631
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-821-234-1631

Query Match 100.0%; Score 412; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.7e-38;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDVCEVCEFLVKEVTKLIDNNKTEKELDAFDKMKSLPKSLSECCQEVVDYSSILS 60
DB 311 SDVCEVCEFLVKEVTKLIDNNKTEKELDAFDKMKSLPKSLSECCQEVVDYSSILS 370
QY 61 ILLEVSPELVCSMLHLCSSG 80
DB 371 ILLEVSPELVCSMLHLCSSG 390

RESULT 2

US-11-078-469-20
; Sequence 20, Application US/11078469
; Publication No. US20050282755A1
; GENERAL INFORMATION:
; APPLICANT: HART, SCOTT A.
; APPLICANT: ZEH, KARIN
; APPLICANT: MACHLEIDT, THOMAS
; APPLICANT: STOLOV, DAVID
; APPLICANT: CONGER, DBE
; TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES THEREOF

us-10-801-517-2.rapbn

Wed Jan 18 11:58:42 2006

```

; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/11/196,400
; CURRENT FILING DATE: 2005-08-04
; PRIOR APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,462
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
; US-11-196-400-3

Query Match 17.0%; Score 70; DB 7; Length 1786;
Best Local Similarity 25.0%; Pred. No. 8.3; Mismatches 20; Indels 4; Gaps 2;
Matches 15; Conservative 21;

Qy 6 EVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILSILLES 65
Db 1238 DITSKLIETQEL---NEVEADLIKMEKL-KELEKALSSEDSKEIIDAKDITLKVIEEE 1293

RESULT 5
US-11-078-469-21
; Sequence 21, Application US/11078469
; Publication No. US20050282755A1
; GENERAL INFORMATION:
; APPLICANT: HART, SCOTT A.
; APPLICANT: ZEH, KARIN
; APPLICANT: MACHLEIDT, THOMAS
; APPLICANT: STOLOV, DAVID
; APPLICANT: CONGER, DEE
; TITLE OF INVENTION: COMPOSITIONS HAVING ANTIMICROBIAL ACTIVITY AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: ANS-2001-UT
; CURRENT APPLICATION NUMBER: US/11/078,469
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 60/554,526
; PRIOR FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: 60/618,948
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 21
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-078-469-21

Query Match 16.4%; Score 67.5; DB 7; Length 145;
Best Local Similarity 21.1%; Pred. No. 0.75; Mismatches 20; Indels 1; Gaps 1;
Matches 15; Conservative 20;

Qy 8 CEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILSILLES 67
Db 69 CLTIVQKLKQWVD-KPTQRSVSNAAITVCGTGRSRWDCVCFNFMRYQSRVTOGLVAGET 127

Qy 68 PELVCSMLHLC 78
Db 128 AQICEDLRLC 138

RESULT 6
US-10-821-234-975
; Sequence 3, Application US/11196400
; Publication No. US20050287166A1
; GENERAL INFORMATION:

```

```
; Sequence 975, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: Pf SEQ_genes Version 1.0
; SEQ ID NO 975
; LENGTH: 989
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-975

Query Match      15.8%; Score 65; DB 6; Length 989;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 20; Conservative 16; Mismatches 31; Indels 10; Gaps 2;

QY 4 YCEVCEPLVKVTKLIDN-NKTEKEILDADFMCCKLPKSLSEBCEQVWDYTGSSI---- 58
Db 592 YCSVIENNKEKAPLFKYQAEIMKLDITLKSQITQASDEADMKAMRMIDELN 651

QY 59 -----LSILLEVSP 70
Db 652 KQVSELSQLYKEAQEL 668

RESULT 7
US-10-878-556A-158
; Sequence 158, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: humangp/chr19-q96g16
; DATABASE ENTRY DATE: 2003-04-22
US-10-878-556A-158

Query Match      15.5%; Score 64; DB 6; Length 476;
Best Local Similarity 31.7%; Pred. No. 7.6;
Matches 19; Conservative 10; Mismatches 19; Indels 12; Gaps 2;

QY 18 LIDNNKTEKETLDA-----FDKMCCKLPKSL--SEBCEQVWDYTGSSILSILLEE 65
Db 244 IVQENKGLKVLIEATKAFLDNPGILSELCTLSRLATRNFCQEVVDLGLSILVSLAD 303

RESULT 8
US-10-467-657-5296
; Sequence 5296, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
```

```
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5296
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5296

Query Match      14.9%; Score 61.5; DB 6; Length 115;
Best Local Similarity 20.5%; Pred. No. 2.6;
Matches 17; Conservative 21; Mismatches 24; Indels 21; Gaps 3;

QY 12 VKEVTKLIDNNKTEKEILDADFMCCKLPKSL---SEEQ-----EVVD----- 52
Db 30 VRDIERLISRNVPRLRIGDAIQNMCMETNVSIWMYRQCKSESSDFVDEMLEHHENPIGS 89

QY 53 --TYGSSILSILLEVSP 73
Db 90 SKSFASLLPLSVTNFTDMACT 112

RESULT 9
US-11-194-246-316
; Sequence 316, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Irepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
; FILE REFERENCE: 00592.US1 (W&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 316
; LENGTH: 292
; TYPE: PRT
; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-316

Query Match      14.8%; Score 61; DB 7; Length 292;
Best Local Similarity 29.0%; Pred. No. 9;
Matches 20; Conservative 9; Mismatches 20; Indels 20; Gaps 3;

QY 12 VKEVTKL-----IDNKT-----EKEILDADFMCCKLPKSLSEBCEQV----- 51
Db 43 IASVTKLTANVFLNNKPNCRITATKEDTDRIKGTGTLKPKNIPISCNELLKAMLVHS 102

QY 52 DTYGSSILS 60
Db 103 DNYAAHALS 111

RESULT 10
US-11-124-368A-276
; Sequence 276, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
```

```
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 2036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-276
```

```
Query Match 14.6%; Score 60; DB 7; Length 2036;
Best Local Similarity 25.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 16; Mismatches 25; Indels 18; Gaps 3;
```

```
QY 17 KLIDNNKTEKILDAFDKMC-S-KLPKSLSEE-----CQEVVDYTGSSILSI 61
Db 281 KLDSAPAEAREPEWKEVCTIELRNGASQRPPLHLCNVDPDLITILHGISETYD---VSP 337
QY 62 LLEEVSPELVCSMLHLCSG 80
Db 338 LLHYMLPHLVVSIHHVTG 356
```

```
RESULT 11
US-11-124-368A-280
; Sequence 280, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 280
; LENGTH: 2036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-280
```

```
Query Match 14.6%; Score 60; DB 7; Length 2036;
Best Local Similarity 25.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 16; Mismatches 25; Indels 18; Gaps 3;
```

```
QY 17 KLIDNNKTEKILDAFDKMC-S-KLPKSLSEE-----CQEVVDYTGSSILSI 61
Db 281 KLDSAPAEAREPEWKEVCTIELRNGASQRPPLHLCNVDPDLITILHGISETYD---VSP 337
QY 62 LLEEVSPELVCSMLHLCSG 80
Db 338 LLHYMLPHLVVSIHHVTG 356
```

```
RESULT 12
US-11-124-368A-281
; Sequence 281, Application US/11124368A
; Publication No. US20050287559A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281
; LENGTH: 2036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-281
```

```
Query Match 14.6%; Score 60; DB 7; Length 2036;
Best Local Similarity 25.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 16; Mismatches 25; Indels 18; Gaps 3;
```

```
QY 17 KLIDNNKTEKILDAFDKMC-S-KLPKSLSEE-----CQEVVDYTGSSILSI 61
Db 281 KLDSAPAEAREPEWKEVCTIELRNGASQRPPLHLCNVDPDLITILHGISETYD---VSP 337
QY 62 LLEEVSPELVCSMLHLCSG 80
Db 338 LLHYMLPHLVVSIHHVTG 356
```

```
RESULT 13
US-10-821-234-916
; Sequence 916, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_SEQ_genes Version 1.0
; SEQ ID NO 916
; LENGTH: 3002
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-916
```

```
Query Match 14.4%; Score 59.5; DB 6; Length 3002;
Best Local Similarity 27.8%; Pred. No. 2.2e+02;
Matches 25; Conservative 8; Mismatches 30; Indels 27; Gaps 5;
```

```
QY 8 CBFL-----VKEVTKLIDNNKTEKEI-LDAFDKMC-S-KLPKSLSEECSQ- 49
Db 2627 CQFLCVNTIGGTCPCPFQTHHTSCIDNNECTSDINLCGSKGICQNTFGSFTCECQRG 2686
QY 50 -VVDYTGSSILSILLEEVSPELVCSMLHLC 78
Db 2687 FSLDQTGSS-----CEDVDE---CEGNHRC 2708
```

```
RESULT 14
US-09-940-308-3
; Sequence 3, Application US/09940308
```

```
; Publication No. US20040092437A1
; GENERAL INFORMATION:
; APPLICANT: Murgita, Robert A.
; TITLE OF INVENTION: Recombinant Human Alpha-Fetoprotein as
; TITLE OF INVENTION: an Immunosuppressive Agent
; FILE REFERENCE: 06727/005003
; CURRENT APPLICATION NUMBER: US/09/940,308
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 09/186,723
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 08/377,309
; PRIOR FILING DATE: 1995-01-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-940-308-3

Query Match      14.3%; Score 59; DB 5; Length 197;
Best Local Similarity 35.1%; Pred. No. 9.3;
Matches 20; Conservative 10; Mismatches 21; Indels 6; Gaps 3;

Qy      7 VCEFLVKEVTKLIDNNKT--EKEILDADFDMC--SKLPKSLSEC--QEVVDYTGSS 57
Db      36 VQEATYKEVSKWVKDALTATEKPTGDSQSGCLENQLPAFLFELCHEKEILEKYGHS 92

RESULT 15
US-09-940-308-6
; Sequence 6, Application US/09940308
; Publication No. US20040092437A1
; GENERAL INFORMATION:
; APPLICANT: Murgita, Robert A.
; TITLE OF INVENTION: Recombinant Human Alpha-Fetoprotein as
; TITLE OF INVENTION: an Immunosuppressive Agent
; FILE REFERENCE: 06727/005003
; CURRENT APPLICATION NUMBER: US/09/940,308
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 09/186,723
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 08/377,309
; PRIOR FILING DATE: 1995-01-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-940-308-6

Query Match      14.3%; Score 59; DB 5; Length 389;
Best Local Similarity 35.1%; Pred. No. 21;
Matches 20; Conservative 10; Mismatches 21; Indels 6; Gaps 3;

Qy      7 VCEFLVKEVTKLIDNNKT--EKEILDADFDMC--SKLPKSLSEC--QEVVDYTGSS 57
Db      36 VQEATYKEVSKWVKDALTATEKPTGDSQSGCLENQLPAFLFELCHEKEILEKYGHS 92

Search completed: January 13, 2006, 16:39:05
Job time : 1.5894 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:26:00 ; Search time 8.34437 Seconds
(without alignments)
792.637 Million cell updates/sec

Title: US-10-801-517-2

Perfect score: 412

Sequence: 1 SDVCEVCEPLVKEVTKLID.....ILLEEVSPNLCVSMHLCSG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCRUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412	100.0	80	1	US-08-100-247-3
2	412	100.0	80	1	US-08-483-146A-3
3	412	100.0	80	1	US-08-232-513A-4
4	412	100.0	80	1	US-08-484-594A-3
5	412	100.0	80	2	US-08-076-258A-3
6	412	100.0	80	2	US-08-756-031-3
7	412	100.0	80	2	US-08-928-074-24
8	412	100.0	523	1	US-08-100-247-2
9	412	100.0	523	1	US-08-483-146A-2
10	412	100.0	523	1	US-08-232-513A-3
11	412	100.0	523	1	US-08-484-594A-2
12	412	100.0	523	2	US-08-076-258A-2
13	412	100.0	523	2	US-08-756-031-2
14	412	100.0	524	2	US-09-352-548-1
15	412	100.0	524	2	US-09-949-016-6272
16	412	100.0	524	2	US-08-928-074-23
17	412	100.0	535	2	US-09-949-016-8603
18	211	51.2	40	2	US-09-780-438C-1
19	201	48.8	38	2	US-09-780-438C-2
20	111	26.9	22	1	US-08-100-247-1
21	111	26.9	22	1	US-08-483-146A-1
22	111	26.9	22	1	US-08-232-513A-1
23	111	26.9	22	1	US-08-484-594A-1
24	111	26.9	22	2	US-09-231-159-1
25	111	26.9	22	2	US-08-611-307-1
26	111	26.9	22	2	US-09-148-030-1
27	111	26.9	22	2	US-09-076-258A-1

28	111	26.9	22	2	US-08-756-031-1	Sequence 1, Appli
29	111	26.9	22	2	US-08-928-074-1	Sequence 1, Appli
30	106	25.7	22	2	US-09-231-159-8	Sequence 8, Appli
31	106	25.7	22	2	US-08-611-307-8	Sequence 8, Appli
32	106	25.7	22	2	US-08-928-074-8	Sequence 8, Appli
33	101	24.5	38	2	US-09-780-438C-3	Sequence 3, Appli
34	101	24.5	38	2	US-09-780-438C-5	Sequence 5, Appli
35	101	24.5	38	2	US-09-780-438C-6	Sequence 6, Appli
36	101	24.5	39	2	US-09-780-438C-4	Sequence 4, Appli
37	96	23.3	22	1	US-08-483-146A-9	Sequence 9, Appli
38	96	23.3	22	1	US-08-232-513A-18	Sequence 18, Appli
39	96	23.3	22	1	US-08-484-594A-9	Sequence 9, Appli
40	96	23.3	22	2	US-09-231-159-7	Sequence 7, Appli
41	96	23.3	22	2	US-08-611-307-7	Sequence 7, Appli
42	96	23.3	22	2	US-09-076-258A-9	Sequence 9, Appli
43	96	23.3	22	2	US-08-928-074-7	Sequence 7, Appli
44	96	23.3	514	2	US-10-339-351-2	Sequence 2, Appli
45	92	22.3	79	2	US-08-596-684F-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-100-247-3
; Sequence 3, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOIN AS A NEUTROTROPIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: OBRIEN.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: SAPOIN C
; US-08-100-247-3

Query Match 100.0%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 6.9e-42;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
QY 61 ILLEEVSPELVCSMLHLCSG 80
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 2
US-08-483-146A-3
; Sequence 3, Application US/08483146A
; Patent No. 5696080
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,146A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-483-146A-3
Query Match 100.0%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 6.9e-42;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
QY 61 ILLEEVSPELVCSMLHLCSG 80
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 3
US-08-232-513A-4
; Sequence 4, Application US/08232513A
```

```
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; TITLE OF INVENTION: as Therapeutic Agents
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..80
; OTHER INFORMATION: /label= Saposin_C
US-08-232-513A-4
Query Match 100.0%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 6.9e-42;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQEVWDTYGSILS 60
QY 61 ILLEEVSPELVCSMLHLCSG 80
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 4
US-08-484-594A-3
; Sequence 3, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
```


QY	1	SDVYCEVCEFLVKEVT	KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS	60
Db	1	SDVYCEVCEFLVKEVT <th>KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS</th> <td>60</td>	KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS	60
QY	61	ILLEVSPVLVCSMLHLC	SG 80	
Db	61	ILLEVSPVLVCSMLHLC	SG 80	
<p>RESULT 8</p> <p>US-08-100-247-2</p> <p>Query Match 100.0%; Score 412; DB 2; Length 80;</p> <p>Best Local Similarity 100.0%; Pred. No. 6.9e-42; Indels 0; Gaps 0;</p> <p>Matches 80; Conservative 0; Mismatches 0;</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: O'BRIEN, JOHN S.</p> <p>APPLICANT: KISHIMOTO, YASUO</p> <p>TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR</p> <p>NUMBER OF SEQUENCES: 5</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR</p> <p>STREET: 620 NEWPORT CENTER DRIVE</p> <p>CITY: NEWPORT BEACH</p> <p>STATE: CA</p> <p>COUNTRY: USA</p> <p>ZIP: 92660</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/100,247</p> <p>FILING DATE: 19930730</p> <p>CLASSIFICATION: 514</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Israel, Ned A.</p> <p>REGISTRATION NUMBER: 29,655</p> <p>REFERENCE/DOCKET NUMBER: O'BRIEN.002A</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 619-235-8550</p> <p>TELEFAX: 619-235-0176</p> <p>INFORMATION FOR SEQ ID NO: 2:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 523 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>HYPOTHETICAL: NO</p> <p>ANTI-SENSE: NO</p> <p>FRAGMENT TYPE: N-terminal</p> <p>IMMEDIATE SOURCE:</p> <p>CLONE: PROSAPOSIN</p> <p>US-08-100-247-2</p>				
QY	1	SDVYCEVCEFLVKEVT <th>KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS</th> <td>60</td>	KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS	60
Db	310	SDVYCEVCEFLVKEVT <th>KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS</th> <td>360</td>	KLIDNNKTEKILDAFDRKMSKLPKSLSEECQEVVDYTGSSILS	360
QY	61	ILLEVSPVLVCSMLHLC	SG 80	
Db	370	ILLEVSPVLVCSMLHLC	SG 389	
<p>RESULT 9</p> <p>US-08-483-146A-2</p> <p>Query Match 100.0%; Score 412; DB 1; Length 523;</p> <p>Best Local Similarity 100.0%; Pred. No. 8e-41; Indels 0; Gaps 0;</p> <p>Matches 80; Conservative 0; Mismatches 0;</p>				

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-09-076-258A-2

Query Match 100.0%; Score 412; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVYGGSSILS 60
DB 310 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVYGGSSILS 369

QY 61 ILLEEVSPELVCSMLHLCSG 80
DB 370 ILLEEVSPELVCSMLHLCSG 389

RESULT 13
US-08-756-031-2
Sequence 2, Application US/08756031
Patent No. 6590074
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,031
FILING DATE: 26-NOV-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'BRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-756-031-2

Wed Jan 18 11:58:42 2006

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-484-594A-2

Query Match 100.0%; Score 412; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVYGGSSILS 60
DB 310 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVYGGSSILS 369

QY 61 ILLEEVSPELVCSMLHLCSG 80
DB 370 ILLEEVSPELVCSMLHLCSG 389

RESULT 12
US-09-076-258A-2
Sequence 2, Application US/09076258A
Patent No. 6559124
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,258A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,970
FILING DATE: 28-OCT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,146
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.2DVLC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176

```
Query Match      100.0%; Score 412; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
    |||||
Db 310 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 369
    |||||

Qy 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 370 ILLEEVSPELVCSMLHLCSG 389
    |||||

RESULT 14
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1

Query Match      100.0%; Score 412; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
    |||||
Db 311 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370
    |||||

Qy 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 371 ILLEEVSPELVCSMLHLCSG 390
    |||||

RESULT 15
US-09-949-016-6272
; Sequence 6272, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6272
```

```
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6272

Query Match      100.0%; Score 412; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 60
    |||||
Db 311 SDVYCEVCEFLVKEVTGLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDYTGSSILS 370
    |||||

Qy 61 ILLEEVSPELVCSMLHLCSG 80
    |||||
Db 371 ILLEEVSPELVCSMLHLCSG 390
    |||||

Search completed: January 13, 2006, 16:36:34
Job time : 8.34437 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:23:49 ; Search time 6.35762 Seconds
(without alignments)
1210.728 Million cell updates/sec

Title: US-10-801-517-2

Perfect score: 412

Sequence: 1 SDVYCVCEPLVKEVTKLID.....ILLERVSPELVCSMLHLCSG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412	100.0	527	1 SAHUP	saposin precursor
2	343	83.3	80	2 S21770	saposin-C - bovine
3	261	63.3	81	2 A32026	glucosylceramide b
4	193	46.8	554	1 A28716	saposin precursor
5	168	40.8	557	1 JH0604	saposin precursor
6	158	38.3	965	2 T00207	p109 protein - sil
7	106.5	25.8	217	2 T48201	hypothetical prote
8	106	25.7	213	2 T46069	hypothetical prote
9	106	25.7	428	2 S47096	cynarase (EC 3.4.2
10	106	25.7	474	2 T12049	cyprosin (EC 3.4.2
11	104	25.2	506	2 F86253	hypothetical prote
12	103	25.0	292	2 T14446	aspartic proteinas
13	103	25.0	322	2 S41400	aspartic proteinas
14	103	25.0	433	2 E96649	hypothetical prote
15	102.5	24.9	376	2 S02766	pulmonary surfacta
16	102	24.8	506	2 T07915	probable aspartic
17	102	24.8	509	2 JC7272	aspartic proteinas
18	100	24.3	513	2 T11686	aspartic proteinas
19	99.5	24.2	103	2 D89587	protein T08A9.8 [i
20	99.5	24.2	363	2 A29072	pulmonary surfacta
21	99	24.0	575	1 A40292	acyloxacyl hydrol
22	98	23.8	496	2 J50732	aspartic proteinas
23	97	23.5	508	2 S19697	aspartic proteinas
24	97	23.5	509	2 S66516	oryzasin (EC 3.4.2
25	97	23.5	513	2 T09739	aspartic endopepti
26	96.5	23.4	101	1 S49145	amoebapore C precu
27	95	23.1	79	1 LNP61	pulmonary surfacta
28	95	23.1	369	2 I46531	surfactant protein
29	92	22.3	509	2 S49349	cyprosin (EC 3.4.2

ALIGNMENTS

RESULT 1

SAHUP

saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protsin; co-beta-glucosidase; componen
ein (SAP); sphingolipid activator protein A2; sulfate activator protein
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence_revision 17-Nov-1995 #text change 09-Jul-2004

C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S3
0226; I37265; I37264

R;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

J. Biochem. 105, 152-154, 1989

A;Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucos
A;Reference number: JX0061; MUID:89255151; PMID:2498298

A;Accession: JX0061

A;Molecule type: mRNA

A;Residues: 1-527 <NAK>

A;Cross-references: UNIPROT:P07602; UNIPARC:UPI000002B33D; GB:D00422; NID:g220063; PIDN
A;Note: alternative splice form 1

A;Accession: A57368

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <NA2>

A;Cross-references: UNIPARC:UPI0000000DBF; GB:J03015; GB:J03086; NID:g337755; PIDN:AAB5
A;Note: alternative splice form 2

R;Rorman, E.G.; Scheink, V.; Grabowski, G.A.

Genomics 13, 312-318, 1992

A;Title: Structure and evolution of the human prosaposin chromosomal gene.

A;Reference number: A42003; MUID:92307663; PMID:1612590

A;Accession: A42003

A;Molecule type: DNA

A;Residues: 50-140 <ROR>

A;Cross-references: UNIPARC:UPI00001741AA; GB:M86181

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107236)

A;Accession: B42003

A;Molecule type: DNA

A;Residues: 185-259;263-276 <RO2>

A;Cross-references: UNIPARC:UPI00001741AB; UNIPARC:UPI00001741AC

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107237)

A;Accession: C42003

A;Molecule type: DNA

A;Residues: 305-393 <RO3>

A;Cross-references: UNIPARC:UPI00001741AD
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107238); sequence in

A;Accession: D42003

A;Molecule type: DNA

A;Residues: 399-487 <RO4>

A;Cross-references: UNIPARC:UPI00001741AE

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107239); sequence in

R;Rorman, E.G.; Grabowski, G.A.

Genomics 5, 486-492, 1989

A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphi
A;Reference number: A30367; MUID:90129043; PMID:2515150

A;Accession: A30367
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <ROS>
A;Cross-references: UNIPARC:UPI0000000DBF; GB:J03077; NID:g183230; PIDN:AAA52560.1; PID:
A;Note: alternative splice form 2
R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdzicka, M.; Fluharty, A.L.; Ginns, E.I.;
Arch. Biochem. Biophys. 304, 110-116, 1993
A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor
A;Reference number: S34740; MUID:93311991; PMID:8323276
A;Accession: S34740
A;Molecule type: protein
A;Residues: 17-24;165-172;180-189;301-305 <HIR>
A;Cross-references: UNIPARC:UPI00001741AF; UNIPARC:UPI00001741B0; UNIPARC:UPI00001741B1;
R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M.
FEBS Lett. 330, 8-12, 1993
A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
A;Reference number: S36140; MUID:93380576; PMID:8370464
A;Accession: S36140
A;Molecule type: protein
A;Residues: 'XX',62,'X',64-65,'X',67-79,'X',81-84 <TY>
A;Cross-references: UNIPARC:UPI00001741B3
A;Note: saposin A
A;Accession: S36141
A;Molecule type: protein
A;Residues: 'XX',413-414,'X',416-428,'X',430-434 <TY2>
A;Cross-references: UNIPARC:UPI00001741B4
A;Note: saposin D
R;Holschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K.
J. Biol. Chem. 266, 7556-7560, 1991
A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and
A;Reference number: S36988; MUID:91210267; PMID:2019586
A;Accession: S36988
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240,'S',242-527 <HOL>
A;Cross-references: UNIPARC:UPI000016AFP6; EMBL:M60255; NID:g337759; PIDN:AAA36594.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative sp
A;Accession: S36989
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240,'S',242-259,263-527 <HO2>
A;Cross-references: UNIPARC:UPI000016AFP7; EMBL:M60257; NID:g337764; PIDN:AAA36595.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative sp
A;Accession: S36990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240,'S',242-259,261-527 <HO3>
A;Cross-references: UNIPARC:UPI000016AFP8; EMBL:M60258; NID:g337766; PIDN:AAA36596.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp
R;Kondoh, K.; Hineno, T.; Sano, A.; Kakimoto, Y.
Biochem. Biophys. Res. Commun. 181, 286-292, 1991
A;Title: Isolation and characterization of prosaposin from human milk.
A;Reference number: PS0330; MUID:92068206; PMID:1958198
A;Accession: PS0330
A;Molecule type: protein
A;Residues: 17-24,'X',26 <KON>
A;Cross-references: UNIPARC:UPI00001741B5
A;Experimental source: milk
R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S.
Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosy
A;Reference number: A35985; MUID:90207231; PMID:2320574
A;Accession: A35985
A;Molecule type: mRNA
A;Residues: 213-221 <KRE>
A;Cross-references: UNIPARC:UPI00001741B6; GB:M32221
A;Accession: B35985
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <KR2>

A;Cross-references: UNIPARC:UPI0000000DBF; GB:M32221; NID:g337761; PIDN:AAA60303.1; PID:
A;Experimental source: lymphoblast
A;Accession: C35985
A;Molecule type: mRNA
A;Residues: 213-216,'I',218-221 <KR3>
A;Cross-references: UNIPARC:UPI00001741B7
A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th
R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.
Eur. J. Biochem. 192, 709-714, 1990
A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein an
A;Reference number: S13195; MUID:91006165; PMID:2209618
A;Accession: S13196
A;Molecule type: protein
A;Residues: 195-259,263-277 <FUE>
A;Cross-references: UNIPARC:UPI00001741B8
R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y.
Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
A;Title: Saposin A: second cerebroside activator protein.
A;Reference number: A32784; MUID:89240739; PMID:2717620
A;Accession: A32784
A;Molecule type: protein
A;Residues: 60-84;86-107;109-119;125-134 <MOR>
A;Cross-references: UNIPARC:UPI00001741B9; UNIPARC:UPI00001741BA; UNIPARC:UPI00001741BB;
R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L.
Science 241, 1098-1101, 1988
A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic
A;Reference number: A41240; MUID:80321660; PMID:2842863
A;Accession: A41240
A;Molecule type: mRNA
A;Residues: 'GSSR',18-259,263-299,'D',301-302,'D',304-527 <OAB>
A;Cross-references: UNIPARC:UPI00001741BD; GB:J03086
R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S.
Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 f
A;Reference number: S02289; MUID:88068647; PMID:2825202
A;Accession: S02289
A;Status: significant sequence differences
A;Molecule type: mRNA
A;Cross-references: EMBL:J03015
A;Note: this sequence corrected by A41240
A;Note: part of this sequence, including the amino end of the mature protein, was determ
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein
A;Reference number: S02028; MUID:89207118; PMID:3242555
A;Accession: S02028
A;Molecule type: protein
A;Residues: 195-259,263-276 <KLB>
A;Cross-references: UNIPARC:UPI00001741BE
R;Fuerst, W.; Machleidt, W.; Sandhoff, K.
Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
A;Title: The precursor of sulfatide activator protein is processed to three different pr
A;Reference number: S00813; MUID:89000190; PMID:3048308
A;Accession: S00813
A;Molecule type: protein
A;Residues: 410-487 <FU2>
A;Cross-references: UNIPARC:UPI00001741BF
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurri
A;Reference number: S00226; MUID:88163077; PMID:3442600
A;Accession: S00226
A;Molecule type: protein
A;Residues: 314-393 <KL2>
A;Cross-references: UNIPARC:UPI000003EDBE
R;Vaccaro, A.M.; Salvati, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano
J. Biol. Chem. 270, 9953-9960, 1995
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bri
A;Reference number: A57297; MUID:95247790; PMID:7730378
A;Contents: annotation; disulfide bonds; glycosylation
R;Holschmidt, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K.
FEBS Lett. 280, 267-270, 1991
A;Title: The organization of the gene for the human cerebroside sulfate activator prote:

[illegible]

C;Superfamily: oryzaasin; saposin repeat homology

Query Match 25.2%; Score 104; DB 2; Length 506;
Best Local Similarity 45.9%; Pred. No. 0.02;
Matches 17; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 43 LSECCQVVDVTGSSILSILLREVSPELVCSMLHLCS 79
 :::: ||| ||||| :||| :||| :||| :|||
DB 313 VSQQKTVVDVGQTILDLLSETQPKKICSGIGLCT 349

RESULT 12
Tl4446
aspartic proteinase (EC 3.4.23.-) - wild cabbage (fragment)
C;Species: Brassica oleracea (wild cabbage)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: Tl4446
R:Fujikura, Y.
submitted to the EMBL Data Library, March 1994
A;Reference number: Z18095
A;Accession: Tl4446
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-292 <FUJ>
A;Cross-references: UNIPROT-Q43407; UNIPARC:UPI00000AC977; EMBL:X77260; NID:g459425; P:
C;Experimental source: cultivar Alpha Paloma; root tips
A;Superfamily: oryzaasin; saposin repeat homology
C;Keywords: aspartic proteinase; hydrolase

Query Match 25.0%; Score 103; DB 2; Length 292;
Best Local Similarity 47.2%; Pred. No. 0.014;
Matches 17; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 44 SEECQEVDVTGSSILSILLREVSPELVCSMLHLCS 79
 :||| :||| :||| :||| :||| :||| :||| :|||
DB 100 SQCKTVVDVGQTILDLLSETQPKKICSGIGLCT 135

RESULT 13
S41400
aspartic proteinase (EC 3.4.23.-) - wild cabbage (fragment)
N;Alternate names: aspartyl protease
C;Species: Brassica oleracea (wild cabbage)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41400
R:Fujikura, Y.; Karssen, C.M.
submitted to the EMBL Data Library, January 1994
A;Description: Cauliflower cDNA with sequence homology to gastric proteases.
A;Reference number: S41400
A;Accession: S41400
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-322 <FUJ>
A;Cross-references: UNIPROT-Q43407; UNIPARC:UPI0000175D78; EMBL:X77260
C;Comment: The pair of saposin repeat homology domains tagged SAP1 and SAP2 represent a
C;Superfamily: oryzaasin; saposin repeat homology
C;Keywords: aspartic proteinase; hydrolase
F;99-144/Domain: saposin repeat homology #status atypical <SAP1>
F;153-203/Domain: saposin repeat homology #status atypical <SAP2>
F;73/Active site: Asp #status predicted

Query Match 25.0%; Score 103; DB 2; Length 322;
Best Local Similarity 47.2%; Pred. No. 0.016;
Matches 17; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 44 SEECQEVDVTGSSILSILLREVSPELVCSMLHLCS 79
 :||| :||| :||| :||| :||| :||| :||| :|||
DB 100 SQCKTVVDVGQTILDLLSETQPKKICSGIGLCT 135

RESULT 14
E96649
hypothetical protein F19K23.21 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96649
C;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, G.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96649
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-433 <STO>
A;Cross-references: UNIPROT:O04593; UNIPARC:UPI000016DA70; GB:AE005173; NID:G2160151; PI
C;Genetics:
A;Gene: F19K23.21
A;Map position: 1
C;Superfamily: oryzasin; saposin repeat homology

Query Match 25.0%; Score 103; DB 2; Length 433;
Best Local Similarity 45.9%; Pred. No. 0.021;
Matches 17; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 43 LSECEQVVDVYGGSSILLESPELVCSMLHCS 79
DB 333 VSQCKTVVDYQGTILDLLAETQPKKICQIGLCA 369

RESULT 15

S02766
pulmonary surfactant protein B precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S02766
Riemrie, P.A.; Shannon, J.M.; Mason, R.J.; Fisher, J.H.
Biochim. Biophys. Acta 994, 215-221, 1989
A;Title: cDNA and deduced amino acid sequence for the rat hydrophobic pulmonary surfacta
A;Reference number: S02766; MUID:89150284; PMID:2920185
A;Accession: S02766
A;Molecule type: mRNA
A;Residues: 1-376 <EMR>
A;Cross-references: UNIPROT:P22355; UNIPARC:UPI00001327F8; EMBL:X14778; NID:G57284; PIDN
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-190/Domain: propeptide #status predicted <PRO>
F;159-151/Domain: saposin repeat homology <SAP1>
F;190-277/Domain: saposin repeat homology <SAP2>
F;191-269/Product: pulmonary surfactant protein B #status predicted <MAT>
F;286-371/Domain: saposin repeat homology <SAP3>

Query Match 24.9%; Score 102.5; DB 2; Length 376;
Best Local Similarity 28.0%; Pred. No. 0.021;
Matches 21; Conservative 18; Mismatches 35; Indels 1; Gaps 1;

QY 5 CEVCEFLVKEVTKLIDNNKTEKILDAFKMCKLP-KSLSECEQVVDVYGGSSILSILL 63
DB 67 CQCEEDIVHLLTWTKEDAFQDTIRKFLQECEDILPLKLVPRCQVLDVLPVLDYFQ 126
QY 64 EVVSEPLVCSMLHLC 78
DB 127 GQIKPRAICSHVGLC 141

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:16:04 ; Search time 183.053 Seconds
(without alignments)
1257.748 Million cell updates/sec

Title: US-10-801-517-1
Perfect score: 2789
Sequence: 1 MYALFLASLLGALAGPVL.....NTETAQCNAVECHRRHVN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	2 AAR70783	Aar70783 Prosaposin
2	2789	100.0	524	2 AAW5652	Aaw5652 Human pro
3	2789	100.0	524	3 AAY58716	Aay58716 Human pro
4	2789	100.0	524	6 ABU79099	Abu79099 Lip-TAA b
5	2789	100.0	524	6 ABU05200	Abu05200 Human exp
6	2789	100.0	524	6 ABU05207	Abu05207 Human exp
7	2789	100.0	524	6 ABU05203	Abu05203 Human exp
8	2789	100.0	524	6 ABU07340	Abu07340 Human exp
9	2789	100.0	524	6 ABU05216	Abu05216 Human exp
10	2789	100.0	524	6 ABU05202	Abu05202 Human exp
11	2789	100.0	524	6 ABU05214	Abu05214 Human exp
12	2789	100.0	524	6 ABU05215	Abu05215 Human exp
13	2789	100.0	524	6 ABU05199	Abu05199 Human exp
14	2789	100.0	524	6 ABU05212	Abu05212 Human exp
15	2789	100.0	524	6 ABU05213	Abu05213 Human exp
16	2789	100.0	524	6 ABU05205	Abu05205 Human exp
17	2789	100.0	524	7 ADP43340	Adp43340 Superanti
18	2789	100.0	524	7 ADJ69401	Adj69401 Human hea
19	2789	100.0	524	8 ADO08060	Ado08060 Human pol
20	2789	100.0	524	8 ADQ94328	Adq94328 Human Pre
21	2789	100.0	524	8 ABM81149	Abm81149 Tumour-as
22	2789	100.0	524	8 ADS87894	Ads87894 Human pro
23	2789	100.0	524	8 ADU48630	Adu48630 Human pro
24	2789	100.0	524	9 ADW80727	Adw80727 Human pro

25	2789	100.0	524	9 ADX06774	Adx06774 Cyclin-de
26	2789	100.0	524	9 ADY14302	Ady14302 PRO polyP
27	2779	99.6	524	6 ABU05208	Abu05208 Human exp
28	2777.5	99.6	527	4 AAB31915	Aab31915 Amino aci
29	2777.5	99.6	527	5 ABP68602	Abp68602 Human pan
30	2777.5	99.6	527	6 ABU79100	Abu79100 Lip-TAA b
31	2777.5	99.6	527	6 ABU05204	Abu05204 Human exp
32	2777.5	99.6	527	6 ABU05210	Abu05210 Human exp
33	2777.5	99.6	527	7 ADF43341	Adf43341 Superanti
34	2772.5	99.4	523	4 AAB31916	Aab31916 Amino aci
35	2772.5	99.4	523	6 ABU05211	Abu05211 Human exp
36	2768	99.2	526	6 ABU05209	Abu05209 Human exp
37	2767.5	99.2	527	6 ABU05206	Abu05206 Human exp
38	2731	97.9	522	8 ADU24090	Adu24090 Human cys
39	2449.5	87.8	479	6 ABR39442	AbR39442 Human GEN
40	2024.5	72.6	385	6 ABR41750	AbR41750 Human DIT
41	1969	70.6	554	7 ADB85295	AdB85295 Rat tubul
42	1921	68.9	554	5 ABB57102	Abb57102 Mouse isc
43	1293.5	46.4	268	8 ADF29875	AdF29875 Human sec
44	1201	43.1	507	8 ADS87898	Ads87898 Human hyp
45	1201	43.1	521	3 AAY56966	Aay56966 Human SBP

ALIGNMENTS

RESULT 1
AAR70783
ID AAR70783 standard; protein; 524 AA.
XX
AC AAR70783;
XX
DT 25-MAR-2003 (revised)
DT 30-AUG-1995 (first entry)
XX
DE Prosaposin.
XX
KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;
KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;
KW adrenal leukodystrophy; prosaposin.
XX
OS Homo sapiens.
XX
PN WO9503821-A1.
XX
PD 09-FEB-1995.
XX
PF 28-JUL-1994; 94WO-US008453.
XX
PR 30-JUL-1993; 93US-00100247.
PR 21-APR-1994; 94US-00232513.
XX
PA (OBRI/) O'BRIEN J S.
XX
PI O'brien JS, Kishimoto Y;
XX
DR WPI; 1995-082029/11.
DR N-PSDB; AAQ85355.
XX
PT Stimulating neural cell out-growth and myelination - with pro:saposin,
PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating
XX nervous system diseases.
PS Disclosure; Page 30-32; 50pp; English.
CC The peptide given in AAR70773, corresponding to amino acids 8-29 of human
CC saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus
CC sequence was determined by comparing the peptide with hematopoietic and
CC neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were
CC identified in the AB loop of human ciliary neurotrophic factor,
CC interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte
CC inhibitory factor, and in helix C of human interleukin-1-beta and
CC oncostatin-M. Prosaponin (AAR70783) and saposin-C also promoted nerve

CC	cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)
CC	(Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
CC	correct PI field.)
XX	
XX	
SQ	Sequence 524 AA;
	Query Match 100.0%; Score 2789; DB 2; Length 524;
	Best Local Similarity 100.0%; Pred. No. 4.6e-237;
	Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MYALFLASLAGAALAGPVGLGKCTRGSAVWCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB	1 MYALFLASLAGAALAGPVGLGKCTRGSAVWCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY	61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB	61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY	121 IKGMSRFGVCSALNLCESLQKHLAELNHQKLESNKI PELDMTEVVAPFMANIPLLY 180
DB	121 IKGMSRFGVCSALNLCESLQKHLAELNHQKLESNKI PELDMTEVVAPFMANIPLLY 180
QY	181 PQDGRSKPQKNDGVQDCIQMTVDIQTAVRTNSTFQALVEHVKECDRLGPGMADI 240
DB	181 PQDGRSKPQKNDGVQDCIQMTVDIQTAVRTNSTFQALVEHVKECDRLGPGMADI 240
QY	241 CKNYISQYSEIAIQMMHMQPKEICALYGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
DB	241 CKNYISQYSEIAIQMMHMQPKEICALYGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
QY	301 PIKKEHVPAKSDVYCEVCEFLVKVTKLIDNNKTEKILIDAPDKMCKSLPKSLSECOEV 360
DB	301 PIKKEHVPAKSDVYCEVCEFLVKVTKLIDNNKTEKILIDAPDKMCKSLPKSLSECOEV 360
QY	361 VDTYSSSTLSILLESVPELVCSMLHLCSTGTELPALTVHVTPQDKGGFCEVCKKLGVYLD 420
DB	361 VDTYSSSTLSILLESVPELVCSMLHLCSTGTELPALTVHVTPQDKGGFCEVCKKLGVYLD 420
QY	421 RNLEKNSTQKEILAALEKCSFLPPYQKQCDQFVAEYEPVLIELVEMDPSPFVCLKIG 480
DB	421 RNLEKNSTQKEILAALEKCSFLPPYQKQCDQFVAEYEPVLIELVEMDPSPFVCLKIG 480
QY	481 ACPSAHKPLLTGTEKINGSPYWCQNTETAACNAVEHCKRHWVN 524
DB	481 ACPSAHKPLLTGTEKINGSPYWCQNTETAACNAVEHCKRHWVN 524
RESULT 2	
AAW85652	
ID	AAW85652 standard; protein; 524 AA.
AC	AAW85652;
XX	
XX	
DT	19-JUL-1999 (first entry)
XX	
DE	Human prosaposin N-terminal peptide.
XX	
KW	Prosaposin; saposin; prosapides; prosaposin receptor agonists; PRA;
KW	peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
KW	therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;
KW	cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
KW	Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;
KW	congestive heart failure; multiple sclerosis;
KW	acute disseminated inflammatory leukoencephalitis;
KW	progressive multifocal leukoencephalitis; Alzheimer's disease;
KW	Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW	ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
KW	cerebral malaria; HTLV; neuropathy;
KW	inflammatory neurodegenerative disease; toxin-induced liver disease.
XX	
XX	Homo sapiens.
OS	
PN	W09912559-A1.

XX	18-MAR-1999.
PD	
XX	09-SEP-1998; 98WO-US019216.
PF	
XX	09-SEP-1997; 97US-0058352P.
PR	
XX	04-JUN-1998; 98US-0088129P.
XX	(REGC) UNIV CALIFORNIA.
PA	
XX	O'brien JS;
PI	
XX	WPI; 1999-229139/19.
XX	N-PSDB; AAX08488.
DR	
PT	
XX	Use of prosaposin receptor agonist.
PS	Claim 7; Fig 2; 90pp; English.
XX	
CC	Prosaposin is a 70kDa glycoprotein which is proteolytically processed to
CC	generate saposins A, B, C and D, all of which are similar to each other
CC	and have a similar placement of six cysteines, a glycosylation site and
CC	conserved proline residues. Prosaposin, saposin C and prosaposin derived
CC	peptides (prosapides), have therapeutic applications in promoting
CC	recovery after toxic, traumatic, myocardial ischaemic, degenerative and
CC	inherited lesions to the peripheral and central nervous system.
CC	Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-
CC	induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt
CC	dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2,
CC	releasing Bcl-2 and its family members which inhibit caspases, thereby
CC	inhibiting apoptosis. An additional mechanism whereby PRAs inhibit
CC	apoptosis is by blocking activation of JNK, a proapoptotic signaling
CC	component. Within several minutes after binding to the receptor, PRAs
CC	block JNK activation induced by tumor necrosis factor-alpha (TNF alpha).
CC	The activation of JNK by TNF alpha is another well known mechanism for
CC	TNF alpha-induced, as well as other proinflammatory cytokine-induced
CC	apoptosis. The method can be used for inhibiting apoptosis which is
CC	caspase-mediated or induced by a proinflammatory cytokine, for example
CC	TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis
CC	associated with a disorder such as e.g. rheumatoid arthritis, Crohn's
CC	disease, irritable bowel syndrome, asthma, cardiac infarction, congestive
CC	heart failure, multiple sclerosis, acute disseminated inflammatory
CC	leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's
CC	disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
CC	disease, ischemic heart disease, Guillain-Barre disease, traumatic brain
CC	injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral
CC	malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and
CC	toxin-induced liver disease. This 524 N-terminal peptide of prosaposin
CC	also acts as a prosaposin receptor agonist
XX	
SQ	Sequence 524 AA;
	Query Match 100.0%; Score 2789; DB 2; Length 524;
	Best Local Similarity 100.0%; Pred. No. 4.6e-237;
	Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MYALFLASLAGAALAGPVGLGKCTRGSAVWCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB	1 MYALFLASLAGAALAGPVGLGKCTRGSAVWCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY	61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB	61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY	121 IKGMSRFGVCSALNLCESLQKHLAELNHQKLESNKI PELDMTEVVAPFMANIPLLY 180
DB	121 IKGMSRFGVCSALNLCESLQKHLAELNHQKLESNKI PELDMTEVVAPFMANIPLLY 180
QY	181 PQDGRSKPQKNDGVQDCIQMTVDIQTAVRTNSTFQALVEHVKECDRLGPGMADI 240
DB	181 PQDGRSKPQKNDGVQDCIQMTVDIQTAVRTNSTFQALVEHVKECDRLGPGMADI 240
QY	241 CKNYISQYSEIAIQMMHMQPKEICALYGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300

Db 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEMPMQTLVPAKVASKNVIPALELVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCKSLPKSLSEECQEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCKSLPKSLSEECQEV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGFCEVCKKLVGYLD 420
QY 421 RNLEKNSTQKEILAALEKGCFLPDPYQKQCDQFVAEYEPVLIIEILVEMDPSFVCLKIG 480
Db 421 RNLEKNSTQKEILAALEKGCFLPDPYQKQCDQFVAEYEPVLIIEILVEMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 3
AAV58716
ID AAV58716 standard; protein; 524 AA.
XX AC
XX AAY58716;
DT 25-APR-2000 (first entry)
XX Human prosaposin.
DE Prosaposin; saponin B; antiangiogenic; angiogenesis inhibitor;
KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;
KW human; therapy.
XX
XX Homo sapiens.
FH Key
FH 195..275 Location/Qualifiers
FT Protein
FT /note= "mature saposin B"
FT Peptide
FT 195..205 /note= "specifically claimed antiangiogenic peptide of
FT Claim 23"
FT Peptide
FT 196..200 /note= "specifically claimed antiangiogenic peptide of
FT Claim 4"
XX
XX WO200002902-A1.
XX
XX 20-JAN-2000.
XX
XX 12-JUL-1999; 99WO-US015772.
XX
XX 13-JUL-1998; 98US-0092647P.
XX
XX (GILL/) GILL P S.
XX
XX Gill PS;
XX
XX WPI; 2000-171128/15.
XX
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and
XX tumor growth.
XX
XX Disclosure; Page 18; 78pp; English.

CC The present sequence is that of human prosaposin, a precursor of saposin
CC B. The invention is based on the discovery that saposin B, previously
CC known to be involved in the hydrolysis of sphingolipids, has potent
CC antiangiogenic and antitumour activity, and also has antiproliferative
CC and antimigratory activity against endothelial cells. This activity is
CC conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684
CC -715), which can be synthetically prepared and used in vitro or in vivo
CC for the treatment of undesired angiogenesis and tumor growth, especially

CC Kaposi's sarcoma (claimed). The polypeptides can also be used in
CC conjunction with cytotoxic moieties to selectively kill certain cell
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,
CC arteriovenous malformation, nonunion fracture, arthritis and other
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
CC trachoma, vascular adhesions and hypertrophic scars
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGAALAGPVGLKECTRGSAVVCQNVKTASDCGAVKHCLQTVWVKTVKS 60
Db 1 MYALFLASLLGAALAGPVGLKECTRGSAVVCQNVKTASDCGAVKHCLQTVWVKTVKS 60
QY 61 LPDCICKOVVTAAGDMLKDNATEEELVYLEKTCDWLKPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKOVVTAAGDMLKDNATEEELVYLEKTCDWLKPKNMSASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
QY 181 PDGPRSRKPPKQDNGDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGWADI 240
Db 181 PDGPRSRKPPKQDNGDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGWADI 240
QY 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEMPMQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEMPMQTLVPAKVASKNVIPALELVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCKSLPKSLSEECQEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCKSLPKSLSEECQEV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGFCEVCKKLVGYLD 420
QY 421 RNLEKNSTQKEILAALEKGCFLPDPYQKQCDQFVAEYEPVLIIEILVEMDPSFVCLKIG 480
Db 421 RNLEKNSTQKEILAALEKGCFLPDPYQKQCDQFVAEYEPVLIIEILVEMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 4
ABU79099
ID ABU79099 standard; protein; 524 AA.
XX
XX AC ABU79099;
XX
XX 18-JUN-2003 (first entry)
XX
XX Lip-TAA binding protein, Prosaposin.
XX
XX Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;
KW gene therapy; mammalian cell receptor; cyostatic;
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
KW tumouricidal immunocyte; antitumour.
XX
XX Unidentified.
XX
XX US2002177551-A1.
XX
XX 28-NOV-2002.
XX

PF	30-MAY-2001; 2001US-00870759.	QY	1	MYALFLIASLLGAALAGPVLGKCEKTRGSVWCONVKTASDCGAVKHCLQTVWVKPTVKS	60
XX		Db	1	MYALFLIASLLGAALAGPVLGKCEKTRGSVWCONVKTASDCGAVKHCLQTVWVKPTVKS	60
XX	31-MAY-2000; 2000US-0208128P.	QY	61	LPCDICRQVVTAGDMLKDNATEBEEILVLEKTCWMLPKPNMSASCKEIVDSYLPVILDI	120
PA	(TERM/) TERMAN D S.	Db	61	LPCDICRQVVTAGDMLKDNATEBEEILVLEKTCWMLPKPNMSASCKEIVDSYLPVILDI	120
XX	Terman DS;	QY	121	IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPPMANIPLLLY	180
PI	WPI; 2003-361759/34.	Db	121	IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPPMANIPLLLY	180
XX		QY	181	PDGPRSKPQKONGDYCQDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI	240
DR	A mammalian cell receptor, useful in the treatment of cancer by binding	Db	181	PDGPRSKPQKONGDYCQDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI	240
PT	to tumor associated lipids where the binding induces anergy or apoptosis	QY	241	CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIIPALELVE	300
XX	in T cells and antigen presenting cells.	Db	241	CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIIPALELVE	300
PS	Disclosure; Page; 167pp; English.	QY	301	PIKKEHVPKSDVYCEVCEFLVKVTKLIDNNKTEKILLDAFDKMCCKLPSLSEECQEV	360
CC	The invention relates to a mammalian cell receptor, useful in the	Db	301	PIKKEHVPKSDVYCEVCEFLVKVTKLIDNNKTEKILLDAFDKMCCKLPSLSEECQEV	360
CC	treatment of cancer, which binds to tumor associated lipids and induces	QY	361	VDYTGSSSTLTLREVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLVGYLD	420
CC	anergy or apoptosis in the T cells and antigen presenting cells (APCs).	Db	361	VDYTGSSSTLTLREVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLVGYLD	420
CC	Also included are a mammalian cell useful in the treatment of cancer	QY	421	RNLEKSNSTKQELIAALEKSGCSFLPDYKQKQDQFVAEYEPVLIRILVEVMDPSFVCLKIG	480
CC	where the receptor which binds tumor associated lipids and induces	Db	421	RNLEKSNSTKQELIAALEKSGCSFLPDYKQKQDQFVAEYEPVLIRILVEVMDPSFVCLKIG	480
CC	cellular inactivation or death is deleted or functionally deactivated,	QY	481	ACPSAHKPLLGTEKICMGPSYWCQNTETAAQCNAVEHCKRHVWN	524
CC	producing (M1) a tumouricidal immunocyte population in vivo in a mammal	Db	481	ACPSAHKPLLGTEKICMGPSYWCQNTETAAQCNAVEHCKRHVWN	524
CC	(by allowing tumour associated lipids to contact immunocytes in which	QY			
CC	receptors for immunosuppressive fatty acids, ceramides, glycolipids,	Db			
CC	sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,	QY			
CC	sialylated glycans, lipopeptides and proteoglycolipids are inactivated or	Db			
CC	deleted), a construct useful in the treatment of cancer comprising a	QY			
CC	superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell	Db			
CC	useful in the treatment of cancer (where an adaptor protein which	QY			
CC	inhibits T cell activation by tumour associated antigens is deleted or	Db			
CC	functionally deactivated), a composition useful in the treatment of	QY			
CC	cancer (comprising a lipid raft conjugated to a superantigen), producing	Db			
CC	(M2) a tumouricidal immunocyte population ex vivo in a mammal (by	QY			
CC	allowing tumour associated lipids to contact immunocytes, in which	Db			
CC	receptors for the lipids are inactivated or deleted to produce a	QY			
CC	tumouricidal immunocyte population, and administering the tumouricidally	Db			
CC	activated immunocytes to the host), producing (M3) a tumouricidal APC	QY			
CC	population ex vivo in a mammal (by allowing a tumour associated lipid to	Db			
CC	contact APCs, in which receptors for the tumour associated lipids are	QY			
CC	inactivated or deleted to produce a tumouricidally activated population,	Db			
CC	and administering APCs to the host), producing a tumouricidal T cell	QY			
CC	population ex vivo in a mammal) by allowing a tumour associated lipids to	Db			
CC	contact T cells, in which adaptor proteins, which inhibit T cell	QY			
CC	activation by tumour associated antigens, are deleted or functionally	Db			
CC	deactivated to produce a tumouricidal population of T cells, and	QY			
CC	administering the tumouricidally activated T cells to the host, or	Db			
CC	allowing a superantigen-lipid raft to contact T cells ex vivo, and	QY			
CC	administering the tumouricidally activated T cells to the host), treating	Db			
CC	(M5) cancer in a mammal (by administering a lipid binding molecule which	QY			
CC	binds immunosuppressive tumour associated lipids in vivo), producing (M6)	Db			
CC	a tumouricidal T cell population in vivo in a mammal (by allowing a	QY			
CC	tumour associated antigen to contact immunocytes in which adaptor	Db			
CC	proteins which inhibit T cell activation by tumour associated antigens	QY			
CC	are deleted or functionally deactivated) and producing (M7) a	Db			
CC	tumouricidal T cell population ex vivo in a mammal comprising allowing a	QY			
CC	superantigen-lipid raft conjugate to contact immunocytes in vivo. The	Db			
CC	receptors, methods and compositions are useful for treating cancers and	QY			
CC	tumours. Bacterial superantigens are co-administered or administered as	Db			
CC	fusion constructs with anti-tumour proteins or motifs. The present	QY			
CC	sequence represents a tumour antigen or a motif identifying a tumour	Db			
CC	antigen, which can be functionally deactivated in the method of the	QY			
CC	invention. Note: The sequence data for this patent did not form part of	Db			
CC	the printed specification, but was obtained in electronic format from the	QY			
CC	US patent office website at	Db			
CC	"seqdata.uspto.gov/sequence.html?DocID=20020177551"	QY			
XX	Sequence 524 AA;	Db			
SQ	Query Match 100.0%; Score 2789; DB 6; Length 524;	QY			
	Best Local Similarity 100.0%; Pred. No. 4.6e-237;	Db			
	Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	QY			

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1866; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC lymphoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLIASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60
Db 1 MYALFLIASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60
QY 61 LPCDICKOVVTAAGDMLKONATBEEILVLEKTCMDLPKPNMNSASCKEIVDSYLPVLIDI 120
Db 61 LPCDICKOVVTAAGDMLKONATBEEILVLEKTCMDLPKPNMNSASCKEIVDSYLPVLIDI 120
QY 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAFPMANIPILLY 180
Db 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAFPMANIPILLY 180
QY 181 PDGPRSKPQKONGDVQCDCIQWTDIQTAVRTNSTFVQALVEHVKEECDRLPGGNADI 240
Db 181 PDGPRSKPQKONGDVQCDCIQWTDIQTAVRTNSTFVQALVEHVKEECDRLPGGNADI 240
QY 241 CKNYISOYSIAIQMMHMQPKEICALVGFCEVKEMPOTLPAKVASKNVIPALVELVE 300
Db 241 CKNYISOYSIAIQMMHMQPKEICALVGFCEVKEMPOTLPAKVASKNVIPALVELVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECEV 360
QY 361 VDTYGSILLSLEVESEFVCSMLHLCSGTRLPALTVHTVTPQKGGFCFCEVCKLGVYLD 420
Db 361 VDTYGSILLSLEVESEFVCSMLHLCSGTRLPALTVHTVTPQKGGFCFCEVCKLGVYLD 420
QY 421 RNLEKSTKQKILAALESKGSFLPDYPKQKCDQVAVSEFVPLIILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTKQKILAALESKGSFLPDYPKQKCDQVAVSEFVPLIILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLTGTEKICWGSPYWCQNTTETAQCNVAHECKRHVN 524
Db 481 ACPSAHKPLLTGTEKICWGSPYWCQNTTETAQCNVAHECKRHVN 524

RESULT 6
ABU05207
ID ABU05207 standard; protein; 524 AA.
XX

AC ABU05207;
XX 29-JAN-2003 (first entry)
DT Human expressed protein tag (EPT) #1873.
DE
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
OS
XX WO200278524-A2.
PN 10-OCT-2002.
PD
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
PA
XX Chicz RM, Tomlinson AJ, Urban RG;
PI WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1873; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLIASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60
Db 1 MYALFLIASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60
QY 61 LPCDICKOVVTAAGDMLKONATBEEILVLEKTCMDLPKPNMNSASCKEIVDSYLPVLIDI 120
Db 61 LPCDICKOVVTAAGDMLKONATBEEILVLEKTCMDLPKPNMNSASCKEIVDSYLPVLIDI 120
QY 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAFPMANIPILLY 180

Db 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
QY 181 PDGPRSKPQPKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
Db 181 PDGPRSKPQPKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPAELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPAELVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQEV 360
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQEV 360
QY 361 VDTYSSSILSILLBEVSPVCSMLHLCSGTGLPALTVHVTPQKDGGEVCKKLVGYLD 420
Db 361 VDTYSSSILSILLBEVSPVCSMLHLCSGTGLPALTVHVTPQKDGGEVCKKLVGYLD 420
QY 421 RNLEKNSTKQETILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQETILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 7
ABU05203
ID ABU05203 standard; protein; 524 AA.
AC ABU05203;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1869.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1869; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide, the purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237; Indels 0; Gaps 0;
Matches 524; Conservative 0; Mismatches 0;
QY 1 MYALFLASLLGAALAGVPLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Db 1 MYALFLASLLGAALAGVPLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLKTCDWLPKPNMSASCKEIVDSYLPVILI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLKTCDWLPKPNMSASCKEIVDSYLPVILI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
QY 181 PDGPRSKPQPKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
Db 181 PDGPRSKPQPKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPAELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPAELVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQEV 360
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQEV 360
QY 361 VDTYSSSILSILLBEVSPVCSMLHLCSGTGLPALTVHVTPQKDGGEVCKKLVGYLD 420
Db 361 VDTYSSSILSILLBEVSPVCSMLHLCSGTGLPALTVHVTPQKDGGEVCKKLVGYLD 420
QY 421 RNLEKNSTKQETILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQETILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 8
ABU07340
ID ABU07340 standard; protein; 524 AA.
XX
XX AC ABU07340;
XX
XX 29-JAN-2003 (first entry)
DT Human expressed protein tag (EPT) #2041.
XX
DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
XX

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
OS
FN WO200278524-A2.
PD 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 2041; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 2789; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-237;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MYALFLASLIGALAGPVLGLKECTRGSAVQNVKTAASDCCAVKHCLQTWNKPTVKS 60
XX
XX 1 MYALFLASLIGALAGPVLGLKECTRGSAVQNVKTAASDCCAVKHCLQTWNKPTVKS 60
XX
XX 61 LPDCIDKDVVTAAGDMLKONATEBEILVILEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
XX
XX 61 LPDCIDKDVVTAAGDMLKONATEBEILVILEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
XX
XX 121 IKGEMSRPGVCSALNLCESLQKHLAEFLNKHQKLESNKIPELDMTEVAVPFMANIPILLY 180
XX
XX 121 IKGEMSRPGVCSALNLCESLQKHLAEFLNKHQKLESNKIPELDMTEVAVPFMANIPILLY 180
XX
XX 181 PQGPRSKPQKNGDVCDPCIQWVTDIQTAVRTNSTFVALVHVHKECDRLGPGNADI 240
XX
XX 181 PQGPRSKPQKNGDVCDPCIQWVTDIQTAVRTNSTFVALVHVHKECDRLGPGNADI 240
XX
XX 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKVASKNVI PALELVE 300
XX

Db 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKVASKNVI PALELVE 300
QY 301 PIKKEHEVPAKSDVYCEVCEFLVKVYTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQV 360
Db 301 PIKKEHEVPAKSDVYCEVCEFLVKVYTKLIDNNKTEKILDAFDKMCCKLPKSLSECCQV 360
QY 361 VDTYSSSILSILLEVSPVLCVSMHLCSGTLTPALTVHVTQPKDGGPCVCKLVGYLD 420
Db 361 VDTYSSSILSILLEVSPVLCVSMHLCSGTLTPALTVHVTQPKDGGPCVCKLVGYLD 420
QY 421 RNLEKNSTKOEITLAALKEGCSFLPDYQKQCFQVAEYEPVLIETILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKOEITLAALKEGCSFLPDYQKQCFQVAEYEPVLIETILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKICWGPSYWCQNTETAAQCNVAVEHCKRHWYN 524
Db 481 ACPSAHKPLLGTEKICWGPSYWCQNTETAAQCNVAVEHCKRHWYN 524
XX
XX RESULT 9
XX ABU05216
XX ID ABU05216 standard; protein; 524 AA.
XX
XX AC ABU05216;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1882.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1882; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing

CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237; Indels 0; Gaps 0;
Matches 524; Conservative 0; Mismatches 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNDTEBEILVYLEKTCMDLWPKPNNMSCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNDTEBEILVYLEKTCMDLWPKPNNMSCKEIVDSYLPVILDI 120

QY 121 IKGMSRPFVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180
Db 121 IKGMSRPFVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180

QY 181 PQDGRSPKPPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Db 181 PQDGRSPKPPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300

QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECOEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECOEV 360

QY 361 VDTYSSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGYLD 420

QY 421 RNLEKNSTKQETLAALKEGCSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQETLAALKEGCSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480

QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVYN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVYN 524

RESULT 10
ABU05202
ID ABU05202 standard; protein; 524 AA.
AC ABU05202;
XX
XX 29-JAN-2003 (first entry)
XX Human expressed protein tag (EPT) #1868.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX

PF 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
DR
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1868; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237; Indels 0; Gaps 0;
Matches 524; Conservative 0; Mismatches 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNDTEBEILVYLEKTCMDLWPKPNNMSCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNDTEBEILVYLEKTCMDLWPKPNNMSCKEIVDSYLPVILDI 120

QY 121 IKGMSRPFVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180
Db 121 IKGMSRPFVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180

QY 181 PQDGRSPKPPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Db 181 PQDGRSPKPPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300

QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECOEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECOEV 360

QY 361 VDTYSSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGYLD 420

QY 421 RNLEKNSTKQETLAALAEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
 Db |||||||
 QY 421 RNLEKNSTKQETLAALAEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
 Db |||||||
 QY 481 ACPSAHKPLLGTEKCIWGSPSYWQNTTETAACNAVEHCRRHVN 524
 Db |||||||
 QY 481 ACPSAHKPLLGTEKCIWGSPSYWQNTTETAACNAVEHCRRHVN 524
 Db |||||||

RESULT 11

ABU05214
 ID ABU05214 standard; protein; 524 AA.

XX

AC ABU05214;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1880.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.

XX Example 2; SEQ ID NO 1880; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 4.6e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYAFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVNNKPTVKS 60
 Db |||||||
 QY 1 MYAFLLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVNNKPTVKS 60
 Db |||||||
 QY 61 LPCDICKDVVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
 Db |||||||
 QY 61 LPCDICKDVVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
 Db |||||||
 QY 121 IKGEMSRFGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 180
 Db |||||||
 QY 121 IKGEMSRFGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 180
 Db |||||||
 QY 181 PQDGPRSRKQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADI 240
 Db |||||||
 QY 181 PQDGPRSRKQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADI 240
 Db |||||||
 QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
 Db |||||||
 QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
 Db |||||||
 QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTGLIDNNKTEKILLDAFDKMCSEKLPKLSSECEQV 360
 Db |||||||
 QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTGLIDNNKTEKILLDAFDKMCSEKLPKLSSECEQV 360
 Db |||||||
 QY 361 VDTYSSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
 Db |||||||
 QY 361 VDTYSSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
 Db |||||||
 QY 421 RNLEKNSTKQETLAALAEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
 Db |||||||
 QY 421 RNLEKNSTKQETLAALAEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
 Db |||||||
 QY 481 ACPSAHKPLLGTEKCIWGSPSYWQNTTETAACNAVEHCRRHVN 524
 Db |||||||
 QY 481 ACPSAHKPLLGTEKCIWGSPSYWQNTTETAACNAVEHCRRHVN 524
 Db |||||||

RESULT 12

ABU05215

ID ABU05215 standard; protein; 524 AA.

XX AC ABU05215;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1881.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX XX

```
PA (ZYCO-) ZYCOS INC.
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1861; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, or protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 2789; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-237;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
QY 61 LPCDICKDVTAAAGDMLKDNATEEILVYLEKTCDWLPKPNMSACKEIVDSILPVLDI 120
Db 61 LPCDICKDVTAAAGDMLKDNATEEILVYLEKTCDWLPKPNMSACKEIVDSILPVLDI 120
QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
QY 181 PDGPRSKPQKNDGVDQDCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Db 181 PDGPRSKPQKNDGVDQDCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
QY 301 PIKKHEVPKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCSPKLSLSECEQEV 360
Db 301 PIKKHEVPKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCSPKLSLSECEQEV 360
QY 361 VDTYGSISILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
Db 361 VDTYGSISILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
QY 421 RNLEKNSITQBIILAALKCCSFLPDYQKQCFVAEYEPVLEIILVEWMDPSFVCLKIG 480
Db 421 RNLEKNSITQBIILAALKCCSFLPDYQKQCFVAEYEPVLEIILVEWMDPSFVCLKIG 480
QY 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAAQCNVHECHCKRHWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAAQCNVHECHCKRHWN 524
```

```
RESULT 13
ABU05199
ID ABU05199 standard; protein; 524 AA.
XX
XX AC ABU05199;
XX
XX 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1865.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX
XX PA (ZYCO-) ZYCOS INC.
XX
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX
XX DR WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1865; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, or protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 2789; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-237;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
```


QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB |||||
QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB |||||
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180
DB |||||
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180
DB |||||
QY 181 PDGPRSKPQKNDGVDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB |||||
QY 181 PDGPRSKPQKNDGVDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB |||||
QY 241 CKNYISOYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVE 300
DB |||||
QY 241 CKNYISOYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVE 300
DB |||||
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSECEQEV 360
DB |||||
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSECEQEV 360
DB |||||
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVTPQKGGFCEVCKLVGYLD 420
DB |||||
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVTPQKGGFCEVCKLVGYLD 420
DB |||||
QY 421 RNLEKNSTKQETILAALEKGCFLPDYPYQKQCOQFVAEYBPVLEILVEVMDPSFVCLKIG 480
DB |||||
QY 421 RNLEKNSTKQETILAALEKGCFLPDYPYQKQCOQFVAEYBPVLEILVEVMDPSFVCLKIG 480
DB |||||
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
DB |||||
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
DB |||||

RESULT 14
ABU05212
ID ABU05212 standard; protein; 524 AA.
XX AC ABU05212;
XX DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1878.
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOS INC.
XX PI Chicz RM, Tomlinson AJ, Urban RG;
XX DR WPI; 2003-040607/03.
XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

PT leukemia.
XX Example 2; SEQ ID NO 1878; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;
SQ Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGAALAGPVGLKCTRGSVAVQNVKTASDCGAVKHCLQTVNKPVTKS 60
DB |||||
QY 1 MYALFLASLLGAALAGPVGLKCTRGSVAVQNVKTASDCGAVKHCLQTVNKPVTKS 60
DB |||||
QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB |||||
QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB |||||
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180
DB |||||
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180
DB |||||
QY 181 PDGPRSKPQKNDGVDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB |||||
QY 181 PDGPRSKPQKNDGVDVCDQCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB |||||
QY 241 CKNYISOYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVE 300
DB |||||
QY 241 CKNYISOYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVE 300
DB |||||
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSECEQEV 360
DB |||||
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSECEQEV 360
DB |||||
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVTPQKGGFCEVCKLVGYLD 420
DB |||||
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVTPQKGGFCEVCKLVGYLD 420
DB |||||
QY 421 RNLEKNSTKQETILAALEKGCFLPDYPYQKQCOQFVAEYBPVLEILVEVMDPSFVCLKIG 480
DB |||||
QY 421 RNLEKNSTKQETILAALEKGCFLPDYPYQKQCOQFVAEYBPVLEILVEVMDPSFVCLKIG 480
DB |||||
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
DB |||||
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
DB |||||
RESULT 15
ABU05213
ID ABU05213 standard; protein; 524 AA.
XX AC ABU05213;
XX DT 29-JAN-2003 (first entry)
XX

DE Human expressed protein tag (EPT) #1879.
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
XX WO200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCO INC.
XX Chicx RM, Tomlinson AJ, Urban RG;
XX MPI, 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX Example 2; SEQ ID NO 1879; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.6e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLLASLLGALAGPVGLKECTRGSAVNCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLLGALAGPVGLKECTRGSAVNCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPCDICKDVTAAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVTAAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
QY 181 PQDGRSKPQPKONGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240

Search completed: January 13, 2006, 16:30:19
Job time : 186.053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:23:49 ; Search time 41.6424 Seconds
(without alignments)
1210.728 Million cell updates/sec

Title: US-10-801-517-1
Perfect score: 2789
Sequence: 1 MVALFLASLLGALAGPVL.....NTETAAQCNAVEHCKRHVMN 524
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2777.5	99.6	527	1 SAHUP	saposin precursor
2	1996	71.6	554	1 A28716	saposin precursor
3	1937.5	69.5	557	1 JH0604	saposin precursor
4	511	18.3	965	2 T00207	P109 protein - sil
5	389	13.9	79	2 A49475	cerebroside sulfat
6	379	13.6	376	2 S02766	pulmonary surfacta
7	348.5	12.5	381	1 LNHUB	pulmonary surfacta
8	343	12.3	80	2 S21770	saposin-C - bovine
9	310	11.1	363	2 A29072	pulmonary surfacta
10	301.5	10.8	369	2 I46531	surfactant protein
11	261	9.4	81	2 A32026	glucosylceramide b
12	256.5	9.2	370	1 LNRBB	pulmonary surfacta
13	213	7.6	213	2 T46069	hypothetical prote
14	203.5	7.3	217	2 T48201	hypothetical prote
15	176.5	6.3	402	2 T15677	hypothetical prote
16	173.5	6.2	513	2 T09739	aspartic endopepti
17	171.5	6.1	506	2 F86253	hypothetical prote
18	165.5	5.9	506	2 T07915	probable aspartic
19	161.5	5.8	513	2 T11686	aspartic proteinas
20	161	5.8	428	2 S47096	cynarase (EC 3.4.2
21	161	5.8	474	2 T12049	cyprosin (EC 3.4.2
22	158.5	5.7	508	2 S19697	aspartic proteinas
23	156	5.6	314	2 T15674	hypothetical prote
24	155.5	5.6	292	2 T14446	aspartic proteinas
25	155.5	5.6	322	2 S41400	aspartic proteinas
26	154	5.5	496	2 JS0732	aspartic proteinas
27	153.5	5.5	509	2 JC7272	aspartic proteinas
28	145.5	5.2	205	2 B89567	protein T08A9.7 [i
29	143.5	5.1	1175	2 S52417	E-selectin ligand-

ALIGNMENTS

RESULT 1

SAHUP

saposin precursor [validated] - human
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfatide sulfatase activator protein
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 17-Nov-1995 #text change 09-Jul-2004
C:Accession: JX0061; A57368; A42003; B42003; D42003; A30367; S34740; S36140; S36
026; I37265; I37264
R:Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.
J. Biochem. 105, 152-154, 1989
A:Title: Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosi
A:Reference number: JX0061; MUID:89255151; PMID:2498298
A:Accession: JX0061
A:Molecule type: mRNA
A:Residues: 1-527 <NA>
A:Cross-references: UNIPROT:P07602; UNIPARC:UPI000002B33D; GB:D00422; NID:g220063; PIDN
A:Note: alternative splice form 1
A:Accession: A57368
A:Molecule type: mRNA
A:Residues: 1-259,263-527 <NA2>
A:Cross-references: UNIPARC:UPI0000000DBF; GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59
A:Note: alternative splice form 2
R:Roman, E.G.; Scheinker, V.; Grabowski, G.A.
Genomics 13, 312-318, 1992
A:Title: Structure and evolution of the human prosaposin chromosomal gene.
A:Reference number: A42003; MUID:92307663; PMID:1612590
A:Accession: A42003
A:Molecule type: DNA
A:Residues: 50-140 <ROR>
A:Cross-references: UNIPARC:UPI00001741AA; GB:M86181
A:Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBI:107236)
A:Accession: B42003
A:Molecule type: DNA
A:Residues: 185-259;263-276 <R02>
A:Cross-references: UNIPARC:UPI00001741AB; UNIPARC:UPI00001741AC
A:Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBI:107237)
A:Accession: C42003
A:Molecule type: DNA
A:Residues: 305-393 <R03>
A:Cross-references: UNIPARC:UPI00001741AD
A:Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBI:107238); sequence inc
A:Accession: D42003
A:Molecule type: DNA
A:Residues: 399-487 <R04>
A:Cross-references: UNIPARC:UPI00001741AE
A:Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBI:107239); sequence inc
R:Roman, E.G.; Grabowski, G.A.
Genomics 5, 486-492, 1989
A:Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin
A:Reference number: A30367; MUID:90129043; PMID:2515150

cyprosin (EC 3.4.2
oryzasin (EC 3.4.2
gene 11-1 protein
probable aspartic
cag pathogenicity
hypothetical prote
aspartic proteinas
cysteine-rich fibr
hypothetical prote
cag island protein
aspartic proteinas
hypothetical prote
hypothetical prote
probable CHP-rich
hypothetical prote
rsec8 - rat (fragm

Wed Jan 18 11:58:40 2006

A;Accession: A30367
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <ROS>
A;Cross-references: UNIPARC:UPI00000000DBF; GB:J03077; NID:G183230; PIDN:AAA52560.1; PID:
A;Note: alternative splice form 2
R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdsicka, M.; Fluharty, A.L.; Gimms, E.I.;
Arch. Biochem. Biophys. 304, 110-116, 1993
A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor
A;Reference number: S34740; MUID:93311991; PMID:8323276
A;Accession: S34740
A;Molecule type: protein
A;Residues: 17-24;165-172;180-189;301-305 <HIR>
A;Cross-references: UNIPARC:UPI00001741B1; UNIPARC:UPI00001741B0; UNIPARC:UPI00001741B1;
R;Tynnelae, J.; Palmer, D.N.; Baumann, M.; Halia, M.
PEBS Lett. 330, 8-12, 1993
A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
A;Reference number: S36140; MUID:93380576; PMID:8370464
A;Accession: S36140
A;Molecule type: protein
A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TTY>
A;Cross-references: UNIPARC:UPI00001741B3
A;Note: saposin A
A;Accession: S36141
A;Molecule type: protein
A;Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TY2>
A;Cross-references: UNIPARC:UPI00001741B4
A;Note: saposin D
R;Holtschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K.
J. Biol. Chem. 266, 7556-7560, 1991
A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and
A;Reference number: S36988; MUID:91210267; PMID:2019586
A;Accession: S36988
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-527 <HOL>
A;Cross-references: UNIPARC:UPI000016AFP6; EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative sp
A;Accession: S36989
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-259, 263-527 <HO2>
A;Cross-references: UNIPARC:UPI000016AFP7; EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative sp
A;Accession: S36990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>
A;Cross-references: UNIPARC:UPI000016AFP8; EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp
R;Kondoh, K.; Hinano, T.; Sano, A.; Kakimoto, Y.
Biochem. Biophys. Res. Commun. 181, 286-292, 1991
A;Title: Isolation and characterization of prosaposin from human milk.
A;Reference number: PS0330; MUID:92068206; PMID:1958198
A;Accession: PS0330
A;Molecule type: protein
A;Residues: 17-24, 'X', 26 <KON>
A;Cross-references: UNIPARC:UPI00001741B5
A;Experimental source: milk
R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S.
Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosy
A;Reference number: A35985; MUID:90207231; PMID:2320574
A;Accession: A35985
A;Molecule type: mRNA
A;Residues: 213-221 <KRE>
A;Cross-references: UNIPARC:UPI00001741B6; GB:M32221
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <KR2>

A;Cross-references: UNIPARC:UPI00000000DBF; GB:M32221; NID:G337761; PIDN:AAA60303.1; PID:
A;Experimental source: lymphoblast
A;Accession: C35985
A;Molecule type: mRNA
A;Residues: 213-216, 'I', 218-221 <KR3>
A;Cross-references: UNIPARC:UPI00001741B7
A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; thi
R;Puerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.
Eur. J. Biochem. 192, 709-714, 1990
A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and
A;Reference number: S13195; MUID:91006165; PMID:2209618
A;Accession: S13196
A;Molecule type: protein
A;Residues: 195-259, 263-277 <FUE>
A;Cross-references: UNIPARC:UPI00001741B8
R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y.
Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
A;Title: Saposin A: second cerebrosidase activator protein.
A;Reference number: A32784; MUID:89240739; PMID:2717620
A;Accession: A32784
A;Molecule type: protein
A;Residues: 60-84;86-107;109-119;125-134 <MOR>
A;Cross-references: UNIPARC:UPI00001741B9; UNIPARC:UPI00001741BA; UNIPARC:UPI00001741BB;
R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L.
Science 241, 1098-1101, 1988
A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic
A;Reference number: A41240; MUID:88321660; PMID:2842863
A;Accession: A41240
A;Molecule type: mRNA
A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>
A;Cross-references: UNIPARC:UPI00001741BD; GB:J03086
R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S.
Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p
A;Reference number: S02289; MUID:88068647; PMID:2825202
A;Accession: S02289
A;Status: significant sequence differences
A;Molecule type: mRNA
A;Cross-references: EMBL:J03015
A;Note: this sequence corrected by A41240
A;Note: part of this sequence, including the amino end of the mature protein, was deter
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein
A;Reference number: S02028; MUID:89207118; PMID:3242555
A;Accession: S02028
A;Molecule type: protein
A;Residues: 195-259, 263-276 <KLE>
A;Cross-references: UNIPARC:UPI00001741BE
R;Fuerst, W.; Machleidt, W.; Sandhoff, K.
Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
A;Title: The precursor of sulfatide activator protein is processed to three different p
A;Reference number: S00813; MUID:89000190; PMID:3048308
A;Accession: S00813
A;Molecule type: protein
A;Residues: 410-487 <FU2>
A;Cross-references: UNIPARC:UPI00001741BF
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurri
A;Reference number: S00226; MUID:98163077; PMID:3442600
A;Accession: S00226
A;Molecule type: protein
A;Residues: 314-393 <KLG>
A;Cross-references: UNIPARC:UPI000003EDBE
R;Vaccaro, A.M.; Salvioli, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano
J. Biol. Chem. 270, 9953-9960, 1995
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bri
A;Reference number: A57297; MUID:95247790; PMID:7730378
A;Contents: annotation; disulfide bonds; glycosylation
R;Holtschmidt, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K.
FEBS Lett. 280, 267-270, 1991
A;Title: The organization of the gene for the human cerebroside sulfate activator prote

A;Reference number: I37264; MUID:91192146; PMID:2013321

A;Accession: I37265

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 59-125 <RES>

A;Cross-references: UNIPARC:UPI000016A751; EMBL:X57107; NID:g30234; PIDN:CAA40391.1; PID

Query Match 99.6%; Score 2777.5; DB 1; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-175;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 180
DB 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 180
QY 181 PDGPRSKPOPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
DB 181 PDGPRSKPOPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
QY 241 CKNYISOYSEITAIQMMHMQDPKBEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALE 297
DB 241 CKNYISOYSEITAIQMMHMQDPKBEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALE 300
QY 298 LVEPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSRE 357
DB 301 LVEPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSRE 360
QY 358 QEVDTYSSILSILLESVSPBELVCSMLHCSGTRLPALTAVHTOPKDGFCFVCKKLVG 417
DB 361 QEVDTYSSILSILLESVSPBELVCSMLHCSGTRLPALTAVHTOPKDGFCFVCKKLVG 420
QY 418 YLDRLNLEKNTKOEITLAELEKCSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCL 477
DB 421 YLDRLNLEKNTKOEITLAELEKCSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCL 480
QY 478 KIGACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
DB 481 KIGACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 527

RESULT 2

A28716
saposin precursor - rat
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A28716
R;Collard, M.W.; Sylvestre, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat s
A;Reference number: A28716; MUID:89000647; PMID:3048385
A;Accession: A28716
A;Molecule type: mRNA
A;Residues: 1-554 <COL>
A;Cross-references: UNIPROT:P10960; UNIPARC:UPI000013555D; GB:M19936; NID:g206904; PIDN:
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfi
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology

C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-554/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;60-143/Product: saposin A #status predicted <SAPA>
F;189-280/Domain: saposin repeat homology <SAP2>
F;194-273/Product: saposin B #status predicted <SAB1>
F;306-397/Domain: saposin repeat homology <SAP3>
F;310-389/Product: saposin C #status predicted <SAPC>
F;431-522/Domain: saposin repeat homology <SAP4>
F;437-514/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,439-512,442-506,470-481/Diulfide bonds: #status predicted
F;80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;197-270,200-264,229-240,314-387,317-381,345-356/Diulfide bonds: #status predicted

Query Match 71.6%; Score 1996; DB 1; Length 554;

Best Local Similarity 66.3%; Pred. No. 4.2e-124;

Matches 368; Conservative 77; Mismatches 78; Indels 32; Gaps 3;

QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 180
DB 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLILY 179
QY 181 PDGPRSKPOPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
DB 180 PDGPRSKPOPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGVSDI 239
QY 241 CKNYISOYSEITAIQMMHMQDPKBEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALEIVE 300
DB 240 CKNYVDQYSEVAVQMMHMQDPKBEICVWVGFCDVEKVPMTLVPAKVASKNVIPALEITD 299
QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPSLSRECOEV 360
DB 300 PYEQDVIQAQNVPCVQCOLVMKLELIINNATELLIKGLSKACSLPAPATKQCEV 359
QY 361 VDTYSSILSILLESVSPBELVCSMLHCSG-----TRLP----- 394
DB 360 LVTFGPSLLDLVLMHEVNPFLCGVISLCSANPNLVGLTLEQPAIAIVGALPKPEAPPKQPE 419
QY 395 -----ALTVHTVTPKGGFCFVCKKLVGYLDRLNLEKNTKOEITLAELEKCSFLPDYQK 449
DB 420 EPKQSALRAHVPPQKNGGFCFVCKKLVILEHNLKNTKEEITLAELEKCSFLPDYQK 479
QY 450 QCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAHKPLLTGTEKCIWGPSYWCNTETA 509
DB 480 QCDEFVAEYEPVLIEILVEVMDPSFVCKIGVCPKSAKLLGTEKCVWGPYWCQNGSETA 539
QY 510 AQCNVAEHCKRHVWN 524
DB 540 ARCNVAHDCKRHVWN 554

RESULT 3

JH0604

saposin precursor - mouse

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: JH0604
R;Teuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992
A;Title: The primary structure of mouse saposin.
A;Reference number: JH0604; MUID:92272718; PMID:1590788

p109 protein - silkworm
C;Species: Bombyx mori (silkworm)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C;Accession: T00207
R;Tambunan, J.; Chang, P.K.; Li, H.; Natori, M.
Gene 212, 287-293, 1998
A;Title: Molecular cloning of a cDNA encoding a silkworm protein which contains the conserved
A;Reference numbers: 214124, MUID:98278844; PMID:9611271
A;Accession: T00207
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-965 <TAM>
A;Cross-references: UNIPROT:O15997; UNIPARC:UPI000007B54A; EMBL:AB008449; NID:92575864; I:
F:778-870/Domain: saposin repeat homology <SAP3>

Query Match 18.3%; Score 511; DB 2; Length 965;
Best Local Similarity 24.6%; Pred. No. 7.5e-26;
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

QY 19 VLGLKECTGSAVWCQNVKTASDCGAVKHCLQTVWVK---PTVKSLPCDICKDVVTAAGD 75
DB 170 LLGSRCTWGPSTWCSNFSSTGRECNATPHCINRVMSKMTFPEDNDNICQICLDMVKQAND 229
QY 76 MLKDNATEBEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDIKIGMSRPGVCSAL 135
DB 230 QLOSNETQDEIKVEFGSKLPIKPFVARGCKMLADEFVVLEIETLASEWN-POAVCSVA 288
QY 136 NLCS--LQKHLAELNHQKLES-----NKIPELDMTEVAFPMW----- 173
DB 289 GLCNAKIDRLVLYNAQRELKAGCYNCKTVGVVRKKPDETKYEDFLVGLLQVCRNMD 348
QY 174 ---NIPLLYP-----QDG-----PRSKQPKDNGD 196
DB 349 LSDSCSNLIPKYVENILEAVKDLNPEGICHVSGQCSYKPHNHDEFTFPEQMVOYSATDD 408
QY 197 V-QDCIQMTVDIQAVRTNSTFVQALVEHVKECDRLGPGMADICKNYISOYSEIAIQM 255
DB 409 VPCEPCEQLVGHRLDVLVANTTELE-FYKVLQGLCKQTKG-KFDECHLHAEQYFVINVF 466
QY 256 MM-HMQPKETCALVGFCDVEKEMPMQTLV-----PAKVA-----SKNVIP 294
DB 467 LVSDLKPAETCKMTGICGNLTSAPISPLVARELVVQPKLIGABESKIARVPLAKOME 526
QY 295 ALELVE--PIKKHEVPA-KSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPK 351
DB 527 ASAAVSVLPLERMFVAAPQSKAACAFQYFLHYLQVQLSDTRTDEKVKAAVQAEACDALPD 586
QY 352 SLSEBCQWVDYTGSSILSLLEVSPELVCSMLHLCSTGLPALTVHVTPQKGGFCFV 411
DB 587 ALNGECKEFVTVQYGSVAVIALVQEIHPASVCPALQICQTE-EIRRVVNSEKSN--CPL 643
QY 412 CKKLGVYLDRLNENKSTQKIELAALKESCELPDPVQKQDQFVAEYEPVLJEILVEVMD 471
DB 644 CLFAVEQLSEVLKNNRSEBENIRKALDGLCTRLSOKLQSECIDFVDITYSSQLVEMLVADWN 703
QY 472 PSFVCLKIGAC-PSAHKPLLGTEKCI 496
DB 704 AKEICVFLKLCRDQLDHPDLKLTSSII 729

RESULT 5
A49475
cerobroside sulfate activator - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 17-May-1996
C;Accession: A49475
R;Stevens, R.L.; Faulk, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.
Biochemistry 32, 4051-4059, 1993
A;Title: Porcine cerobroside sulfate activator: further structural characterization and
A;Reference number: A49475; MUID:93229506; PMID:8471613
A;Accession: A49475
A;Status: preliminary
A;Molecule type: protein

A;Accession: JH0604
A;Molecule type: mRNA
A;Residues: 1-557 <TSU>
A;Cross-references: UNIPROT:O61207; UNIPARC:UPI0000170C25; GB:S36200; NID:9249386; PIDN:
A;Experimental source: liver
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
C;Keywords: signal sequence #status predicted <SIG>
F;1-16/Domain: saposin A #status predicted <SAP>
F;17-557/Product: saposin repeat homology <SAP1>
F;50-143/Domain: saposin A #status predicted <SAP2>
F;189-283/Product: saposin B #status predicted <SAB1>
F;194-276/Product: saposin B #status predicted <SAP3>
F;309-400/Domain: saposin C #status predicted <SAPC>
F;313-392/Product: saposin C #status predicted <SAP4>
F;434-525/Domain: saposin repeat homology <SAPD>
F;440-517/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4
F;80,214,334,379,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.5%; Score 1937.5; DB 1; Length 557;
Best Local Similarity 63.6%; Pred. No. 3.1e-120;
Matches 355; Conservative 77; Mismatches 91; Indels 35; Gaps 3;

QY 1 MYALFLASLGAALAGVLGLKECTGSAVWCQNVKTASDCGAVKHCLQTVWVKPKS 60
DB 1 MYALALPASLLATALTSPVDPKTCGSGSVLCRLDKTAVDCGAVKHCCQMVWSEKPTAKS 60
QY 61 LPDCIDKDVVTAAGDMLKDNATEBEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPDCIDKTVTEAGNLLKDNATQBEILHYLEKTCBIHDSLSASCKEIVDSYLPVILDM 120
QY 121 IKGMSRPGVCSALNLCSLQKHLAELNHQKLESNKIPELDMTEVAFPMWNPILLY 180
DB 121 IKGMSNPGVCSALNLCSLQKHLAELNHQKLESNKIPELDMTEVAFPMWNPILLY 179
QY 181 PDGPRSKPQKONGDVQDCDQMTVDIQAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 180 PDGPRSQPQPRANEDVCQDKMLVSDVQTAVKTNSSFTQGFVDHVKEDCDRLGPGVSDI 239
QY 241 CKNYISOYSEIAIQMGMH---OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPA 297
DB 240 CKNYVDQYSEVQVLMHMQDQPKICVLAGFCNEVVRPKTLVPATETIKNILPAL 299
QY 298 LVPEPKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPSLSE 357
DB 300 MMDPEQNLVQAHVILCQTCQFVWKFSELI VNNATEELLVGLSNACVGLPDPARTKC 359
QY 358 QEVVDYTGSSILSLLEVSPELVCSMLHLCSG----- 390
DB 360 QEVVGTGFPILLDFIHEVNPSSLCGVIGLCAARPELVLEALEOPAPAIVSALLKEPTPK 419
QY 391 ----TRLPALTVAHVTPQKGGFCVCKKLGVYLDRLNENKSTQKIELAALKESCELPDP 446
DB 420 QPAQPKQSALPAHVPPQKNGGFCVCKKLGVYLDRLNENKSTQKIELAALKESCELPDP 479
QY 447 YQKQCDQFAVEYEPVLJEILVEVMDPSFVCLKIGACPSAHKPLLGTEKCTWGPYWCNT 506
DB 480 YQKQCDQFAVEYEPVLJEILVEVMDPGFVCSKIGVCSAYKLLGLGTEKCVWGSPYWCNM 539
QY 507 ETAAQCNVAVHCKRHVN 524
DB 540 ETAAQCNVAVHCKRHVN 557

RESULT 4
T0207

A;Residues: 1-79 <STE>
A;Cross-references: UNIPARC:UPI0000177938
A;Experimental source: kidney
A;Note: sequence extracted from NCBI backbone (NCBIP:129597)
C;Superfamily: saposin; saposin repeat homology
F;1-79/Domain: saposin repeat homology <SAP>
Query Match 13.9%; Score 389; DB 2; Length 79;
Best Local Similarity 88.6%; Pred. No. 3.6e-19;
Matches 70; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 195 GDVQCDCIQWTDIQAVRTNSTFVQALVHVHKECDRLGPGNADICKNYISQYSIAIQ 254
DB 1 GDVQCDCIQWTDQNAVRTNSTFVEALVNHAKCEDRLGPGNADCKNYISQYSIAIQ 60
QY 255 MWMHMQPKICALVGCDE 273
DB 61 MWMHMQPKDICALVGCDE 79
RESULT 6
S02766
pulmonary surfactant protein B precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S02766
R;Enrie, P.A.; Shannon, J.M.; Mason, R.J.; Fisher, J.H.
Biochim. Biophys. Acta 994, 215-221, 1989
A;Title: cDNA and deduced amino acid sequence for the rat hydrophobic pulmonary surfactant protein B
A;Reference number: S02766; MUID:89150284; PMID:2920185
A;Accession: S02766
A;Molecule type: mRNA
A;Residues: 1-376 <EMR>
A;Cross-references: UNIPROT:P22355; UNIPARC:UPI000013278; EMBL:X14778; MUID:957284; PIDN:1000013278
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-190/Domain: propeptide #status predicted <PRO>
F;59-151/Domain: saposin repeat homology <SAP1>
F;190-277/Domain: saposin repeat homology <SAP2>
F;191-269/Domain: saposin repeat homology <SAP3>
F;286-371/Domain: saposin repeat homology <SAP3>
Query Match 13.6%; Score 379; DB 2; Length 376;
Best Local Similarity 22.6%; Pred. No. 1.2e-17;
Matches 116; Conservative 75; Mismatches 158; Indels 164; Gaps 16;
QY 4 LFLASILGAAAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNPVKSLPC 63
DB 9 LALLPTLCISLGATESASSPDCAQGPKEFCQSLQEQALQCFALGCHLQEVWGHAGANDL-C 67
QY 64 DICQVVTAAAGDMLKONATEBEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIK 123
DB 68 QECEDIVHLLTKMTKEDAFQDTIRKFLQECEDILPLKLLVPRCQVLDVYLPVDFYQ 127
QY 124 EMSPRGVCSALNLCESLQHLAELNHQKLESNKIPELMDTEVWAPPFMANIPLLVLPQD 183
DB 128 QI-KPKAICSHVGLC-----PE-----GQT 146
QY 184 GPRSKPQKQNGDVQCDCIQWTDIQAVRTNSTFVQALVHVHKECDRLGPGNADICKN 243
DB 147 KPEQKPE----- 153
QY 244 YISQYSIAIQMMHMQPKICALVGCDEVKEMPMQTLVPKAVSKNVPALPE-----LVE 300
DB 154 -----MLDAIPENPLNKLVLPLFGAPFLAR 178
QY 301 P-----IKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILD-AFDKMCCKLPKSL 353
DB 179 FGPHTQDLSEQQLEPIPLP-FCWLCTRLIKRQAVI-----PKGVLAIVASQVCHVPLVV 232
QY 354 SEECQEVVDYTGSSILGILLESVPELVCSMLHCS-----GTRLPALTVHVTVQ-PKDG 407
DB 233 GGICQCLAEYTVLLDALLGRVVPQLVCGVLVLECSADAIGPALPALPELIEKWFPLQDT 292

QY 408 FCEVCKKLVGYLDRNLKNSTRKQETLAALKKGC-SFLPPYQKQCDQFVARYEPVLIEL 466
DB 293 ECHFKCKSVI-----NQAWNTSQAMPOAMHQACLRFLWD--RQKCEQFVEQHMPQLLALV 345
QY 467 VEVMPSPFVCLKIGACPSAHKPLLLGTEKCIWGP 499
DB 346 PRSQAHTSCQALGVCEAPASEL-----QCFQTP 374
RESULT 7
LNHUB
pulmonary surfactant protein B precursor [validated] - human
N;Alternate names: pulmonary surfactant proteolipid SP-B; pulmonary surfactant-associate
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: A31361; A28461; A27794; A27592; J00162; S21134
R;Pilot-Matias, T.J.; Kister, S.E.; Fox, J.L.; Kropp, K.; Glasser, S.W.; Whitsett, J.A.
DNA 8, 75-86, 1989
A;Title: Structure and organization of the gene encoding human pulmonary surfactant protein B
A;Reference number: A31361; MUID:89170128; PMID:2924687
A;Accession: A31361
A;Molecule type: DNA
A;Residues: 1-381 <PIL>
A;Cross-references: UNIPROT:P07988; UNIPARC:UPI00001423D4; GB:M24461
A;Note: this protein is encoded by a single gene
R;Jacobs, K.A.; Phelps, D.S.; Steinbrink, R.; Fisch, J.; Kriz, R.; Mitsock, L.; Dougher, J.
J. Biol. Chem. 262, 9808-9811, 1987
A;Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa
A;Reference number: A28461; MUID:87250653; PMID:3597440
A;Accession: A28461
A;Molecule type: mRNA
A;Residues: 1-227, 'A', 229-381 <JAC>
A;Cross-references: UNIPARC:UPI00000162D; GB:J02761; NID:9190673; PIDN:AAA60212.1; PID:100000162
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Glasser, S.W.; Korfagen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987
A;Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associate
A;Reference number: A27794; MUID:87231940; PMID:3035561
A;Accession: A27794
A;Molecule type: mRNA
A;Residues: 'EPR', 99-317, 'L', 319-381 <GLA>
A;Cross-references: UNIPARC:UPI000014237D; GB:M16764; NID:9338410; PIDN:AAA88099.1; PID:1000014237
A;Note: 131-ile was also found
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Cochran, J.
J. Clin. Invest. 81, 826-833, 1988
A;Title: Use of human surfactant low molecular weight apoproteins in the reconstitution of human pulmonary surfactant
A;Reference number: A27592; MUID:88139786; PMID:3343343
A;Accession: A27592
A;Molecule type: mRNA
A;Residues: 139-177, 'V', 179-227, 'A', 228-381 <REV>
A;Cross-references: UNIPARC:UPI00001741A7; GB:M19097
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Johansson, J.; Joernvall, H.; Curstedt, T.
FEBS Lett. 301, 165-167, 1992
A;Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and peptide
A;Reference number: S21134; MUID:92233937; PMID:1568474
A;Accession: S21134
A;Status: preliminary
A;Molecule type: protein
A;Residues: 201-227, 'I', 229-279 <JOH>
A;Cross-references: UNIPARC:UPI00001741A9

A>Note: 228-Ala was also found
C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C:Genetics:
A:Gene: GDB:SFTPB; SFTP3; SP-B
A:Cross-references: GDB:I20374; OMTM:178640
A:Map position: 2p12-2p11.2
A:Introns: 23/1; 65/3; 89/3; 131/3; 194/3; 224/3; 286/1; 334/3; 361/3
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
C:Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-200/Domain: propeptide #status predicted <PRO>
F:61-153/Domain: saposin repeat homology <SAP1>
F:200-287/Domain: saposin repeat homology <SAP2>
F:201-279/Product: pulmonary surfactant protein B, 9K form #status predicted <SP9>
F:201-256/Product: pulmonary surfactant protein B, 6K form #status experimental <SP6>
F:291-376/Domain: saposin repeat homology <SAP3>
F:69-143,72-137,100-112,299-366,302-360,325-335/Disulfide bonds: #status predicted
F:129,311/Binding site: carbohydrate (Aan) (covalent) #status predicted
F:208-277,211-271,235-246/Disulfide bonds: #status experimental
F:248/Disulfide bonds: interchain #status experimental

Query Match 12.5%; Score 348.5; DB 1; Length 381;
Best Local Similarity 21.6%; Pred. No. 1.2e-15;
Matches 111; Conservative 72; Mismatches 171; Indels 161; Gaps 16;

Qy 2 VALFLLASILGALGAPVILGLKECTRGSAVMQNVKTASDCGAVKHCLQTVNNKPTVKS 61
Db 9 WLLLLPLTLCGPTAAWTTSSLACAGPEFWCOSLEALQCRALGHCLQEVGHVGADDL 68
Qy 62 PCDICKDVVTAAGMDKONATEBEIILVYLEKTCOMLPKNMSASCKEIVDSYLPVILDI 121
Db 69 -COECEDIHILNKWAKEAIFQDTMRKFLQECNVLPKLMLPQCQNVLDYFFPLVIDYF 127
Qy 122 KEEMSRPGEVCSALMLCSLQKHLAEHLNQLESNKIPELDMTEVVAFFMANIPLLLYP 181
Db 128 QNQTDSNG-ICMHLGLCKSRQ----- 147
Qy 182 QDGPRSKPQKQDNGVQCDCIQWTDIQTAVRTNSTFQALVEHVKEECDRGLPGWADIC 241
Db 148 -----PEPEQE-----PGMSD-- 158
Qy 242 KNYISQYSBIAIQMMHMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVTPALLEVEP 301
Db 159 -----PLPK-----PLRDLPLDPLDLKLVLP--GA 184
Qy 302 IKKHEVPAKSDV-----YCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPK 351
Db 185 LQARPGPHQTDLSEQQFPPLPYCMLCRALIKRIQMIPIKG-----ALRVAVAQCRVVPL 240
Qy 352 SLSEECQEVVDTVGSSILSILLESVPCLVCSMLHLC-----GTRLPALTVHVTPQKDG 406
Db 241 VAGGICQCLAEYSVILLDTLGRMLPQLVCRVLVLECSMDSAGPRSP---TGEMLPKDS 297
Qy 407 GFCEVCKGLGVLDNRLEKNSTKQETLAALKEGC--SFLPDPYQKQCDQFVAYEYEVLE 464
Db 298 E-CHLCMSVT-----TQAGNSSEQAIPQAMLAQCVGSWLD---REKCKQFVEQHTPQLIT 348
Qy 465 ILVEMDPSFVLKIGACPSAHKPLLGTEKCIWGP 499
Db 349 LVPRGWDHATTTCQALGVCGTMSPL-----QCIHSP 379

RESULT 8
S21770
saposin-C - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S21770
R:Sano, A.; Mizuno, T.; Kondoh, K.; Hineno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.
Biochim. Biophys. Acta 1120, 75-80, 1992
A:Title: Saposin-C from bovine spleen; complete amino acid sequence and relation between
A:Reference number: S21770; MUID:92207994; PMID:1554743
A:Accession: S21770

A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-80 <SAN>
A:Cross-references: UNIPARC:UPI0000177939
C:Superfamily: saposin; saposin repeat homology

Query Match 12.3%; Score 343; DB 2; Length 80;
Best Local Similarity 79.7%; Pred. No. 3.9e-16;
Matches 63; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEVVDTVGSSILS 370
Db 1 ADYQCQCEFWKVEKALIDNNRTEELHALDKVCSKLPSTSLAEQCEQEVVDTVGSRILS 60
Qy 371 ILLEEVSPCLVCSMLHLC 389
Db 61 ILLDEASPELVCSMLHLC 79

RESULT 9
A29072
pulmonary surfactant protein SP 18 precursor - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: B29072; A29072
R:Hawgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.
Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987
A:Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and e
A:Reference number: A29072; MUID:87092398; PMID:3467361
A:Accession: B29072
A:Molecule type: mRNA
A:Residues: 1-363 <HAW>
A:Cross-references: UNIPROT:P17129; UNIPARC:UPI00001327F3; GB:M15170; NID:G164077; PIDN:
A:Accession: A29072
A:Molecule type: protein
A:Residues: 182-210 <HAW>
A:Cross-references: UNIPARC:UPI0000177937
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-180/Domain: propeptide #status predicted <PRP>
F:54-146/Domain: saposin repeat homology <SAP1>
F:180-267/Domain: saposin repeat homology <SAP2>
F:181-363/Product: pulmonary surfactant protein SP 18 #status experimental <MAT>
F:273-358/Domain: saposin repeat homology <SAP3>

Query Match 11.1%; Score 310; DB 2; Length 363;
Best Local Similarity 22.0%; Pred. No. 3.9e-13;
Matches 112; Conservative 76; Mismatches 161; Indels 160; Gaps 20;

Qy 4 LFLASL--LGAA-LAGPVILGLKECTRGSAVMQNVKTASDCGAVKHCLQTVNNKPTVKS 60
Db 4 LLLPLTLCGLGAADWSAPSLA---CARGPAFWCOSLEALQCRALGHCLQEVGNARADD 60
Qy 61 LPCDICKDVVTAAGMDKONATEBEIILVYLEKTCOMLPKNMSASCKEIVDSYLPVILDI 120
Db 61 L-CQECQDVIIVLTQMTKEAIFQDMVRKFLHECDVLPKLTTPQCHEMLGTYFFVVDY 119
Qy 121 IKGEMSRPGEVCSALMLCSLQKHLAEHLNQLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 120 FQSQIN-PKILCKHLGLC---KPGLPPEPEQSELSPLDLKLLPEL----- 162
Qy 181 PQDGPRSKPQKQDNGVQCDCIQWTDIQTAVRTNSTFQALVEHVKEECDRGLPGWADI 240
Db 163 -----PGALQV 168
Qy 241 KKNYISQYSBIAIQMMHMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVTPALLEVE 300
Db 169 TGPHTQDUSE-----QQLP-----IP----- 184
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQ 359
Db 185 -----LP-----YCWLCRTLLIKRIQAMI-----PKGVLAVTGQVCHVVPLVVGICQC 228

[illegible][illegible]

Wed Jan 18 11:58:40 2006

C:Accession: T48201
 R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: Z24488
 A:Accession: T48201
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-217 <BEV>
 A:Cross-references: UNIPROT:Q9LZW6; UNIPARC:UPI00000A6022; EMBL:AL162351
 A:Experimental source: cultivar Columbia; BAC clone T20L15
 C:Genetics:
 A:Map position: 5
 A:Introns: 30/1; 79/3; 146/3; 166/3
 A:Note: T20L15.70

Query Match 7.3%; Score 203.5; DB 2; Length 217;
 Best Local Similarity 27.8%; Pred. No. 2.2e-06;
 Matches 52; Conservative 40; Mismatches 82; Indels 13; Gaps 7;
 298 LVEPIKK-HEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEE 356
 25 LLEPFSAHD-----DNQVCELDKYYTLVDYLDQYDNQNELVEALHISCQIP-PLKQK 79
 357 QQEWDTYSSILSILLESVSPVLVCSMLHLCSTRLPALTVHTVTPKDGFCFVCKKL 416
 80 CLSWDHY-TQLPFTQVSTIKSDQICKRLNLCQAV-TPAFASQVHQ-----GNCEACRETV 133
 417 GYLDRNLEKNSKQETILALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMDPSFVC 476
 134 SEVVTKLDPETKLIIRLLKECKSL-NNYQDKCKWVEYGYPLMLTDLQKFLKKDVC 192
 477 LKIGACP 483
 193 TLHVCP 199

RESULT 15

T15677
 hypothetical protein C28C12.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T15677
 R:Miller, N. The EMBL Data Library, November 1995
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid C28C12.
 A:Reference number: Z18387
 A:Accession: T15677
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-402 <MIL>
 A:Cross-references: UNIPROT:Q18279; UNIPARC:UPI0000076C88; EMBL:U40797; NID:gl065916; P
 A:Experimental source: strain Bristol N2; clone C28C12
 C:Genetics:
 A:Gene: CESP:C28C12.5
 A:Map position: 4
 A:Introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 6.3%; Score 176.5; DB 2; Length 402;
 Best Local Similarity 20.3%; Pred. No. 0.00029;
 Matches 75; Conservative 58; Mismatches 156; Indels 81; Gaps 15;
 184 GPRSKPQKNGDVQCDCIQMTVDIQAVRTNSTFVQALVHVKECDRLGPGHADIEN 243
 16 GAQSAPSP-----CESCKSWQNFIDASKDRMKMAQLKVS-LSMCL--VGTSHQSDCK 66
 244 YISOYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPI 302
 67 TLKLDLFTAYKLAPYLADTSVACSKLQWCGSSQFSPLARLAWLYLKGEAIVA---NDNI 123
 303 KKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECEVVD 362
 124 MRQEV-----CDECOASTAQIGKLVGDEFTTYAVKSTLQRFVCKSAGAKHAC----N 172

131 QINAKA-ICQHLGLCQ-----PGSPEPPLD-----PLPKDLVLPRTL---G 167
 184 GPRSKPQKNGDVQCDCIQMTVDIQAVRTNSTFVQALVHVKECDRLGPGHADIEN 243
 168 ALPAKPGP----- 175
 244 YISOYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPIK 303
 176 -----HTQ-----DLSAQPPFPIPL----- 190
 304 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECEVVD 362
 191 -----LWLCRLTLKRIQAMI-----PKGVLMAVAQVCHVPLVVGIGICQCLAE 235
 363 TYGSSILSILLESVSPVLVCSMLHLCSTRLPALTVHTVTPKDGFCFVCKKL 415
 236 RYTVILDEVILGHVLPOLVGLVLRCSVDSIGQVPTLEALPQEWLPQDPE-CRLCMSV 294
 416 VGYLDRNLEKNSKQETILALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMD 470
 295 TTQA-RNISEQTRPQAVVTHACLSQLDK-----QCEQFVAHAP-----AA 335
 471 DPSFVCLKIGACPSAKHPLGTE-----KCIWGPSY 501
 336 EPAVQGL---GCP---RNLPGFGRVVATLSPLQCIQSPHF 370

RESULT 13

T46069
 hypothetical protein T18N14.110 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T46069
 R:Deisen, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z23013
 A:Accession: T46069
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213
 A:Cross-references: UNIPROT:Q9SCT5; UNIPARC:UPI00000AC273; EMBL:AL1132968
 A:Experimental source: cultivar Columbia; BAC clone T18N14
 C:Genetics:
 A:Map position: 3
 A:Introns: 31/1; 80/3; 146/3; 166/3
 A:Note: T18N14.110

Query Match 7.6%; Score 213; DB 2; Length 213;
 Best Local Similarity 29.7%; Pred. No. 5e-07;
 Matches 57; Conservative 35; Mismatches 90; Indels 10; Gaps 6;
 292 VIPALELVEPIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPK 351
 18 VSDARSFVDTISEKVNKSDV-CTLCEEYVTDALSYLEKNVTQAEIIEIDLHRCSQL-R 75
 352 SLSECEQWVDYTGSSILSILLESVSPVLVCSMLHLCSTRLPALTVHTVTPKDGFCFV 411
 76 GYSQCCLSLVDY-VPLFFQLQESFQPHYFCKRMNLG--KVALVEEARQDS-----CGV 128
 412 CKLGVGLDRNLEKNSKQETILALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMD 471
 129 CHRTVSEILIKLQDDPTQLDIVELLIKGCKSLKN-YEKKCKTLVFEYGFILVNAERFLV 187
 472 PSFVCLKIGACP 483
 188 KNDVCTLLRACP 199

RESULT 14

T48201
 hypothetical protein T20L15.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:22:54 ; Search time 199.536 Seconds
(without alignments)
1852.779 Million cell updates/sec

Title: US-10-801-517-1

Perfect score: 2789

Sequence: 1 MYALFILASLLGALAGPYL.....NTTAAQCNAVERCKRHVWN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2789	100.0	524	1 SAP HUMAN	P07602 h proactiva
2	2789	100.0	524	2 Q5JQ36 HUMAN	Q5JQ36 homo sapien
3	2785	99.9	524	2 Q53R35 HUMAN	Q53R35 homo sapien
4	2783	99.8	524	2 Q5JQ37 HUMAN	Q5JQ37 homo sapien
5	2777.5	99.6	530	2 Q59EN5 HUMAN	Q59EN5 homo sapien
6	2776.5	99.6	527	2 Q5NVDS POMPY	Q5NVDS pongo pygma
7	2770	99.3	526	2 Q5R4U7 POMPY	Q5R4U7 pongo pygma
8	2451.5	87.9	497	2 Q4R590 MACFA	Q4R590 m brain cdn
9	2425.5	87.0	525	1 SAP BOVIN	P26779 b proactiva
10	2374.5	85.1	452	2 Q5R406 POMPY	Q5R406 pongo pygma
11	1996	71.6	554	1 SAP RAT	P10960 rattus norv
12	1982.5	71.1	553	2 Q6P7A4 RAT	Q6P7A4 rattus norv
13	1950	69.9	554	2 Q8BFQ1 MOUSE	Q8BFQ1 m mus muscu
14	1938.5	69.5	557	1 SAP MOUSE	Q61207 mus musculus
15	1679.5	60.2	518	1 SAP CHICK	O13035 gallus gall
16	1442	51.7	512	2 Q7S770 XENLA	O7S770 xenopus lae
17	1429	51.2	518	2 Q64286 XENLA	Q64286 xenopus lae
18	1416	50.8	550	2 Q4RQ38 TETNG	Q4RQ38 tetraodon n
19	1414	50.7	520	2 Q8UVZ4 BRARE	Q8UVZ4 brachydanio
20	1412	50.6	520	2 Q6PH48 BRARE	Q6PH48 brachydanio
21	1411	50.6	520	2 Q6PG77 BRARE	Q6PG77 brachydanio
22	1343	48.2	522	2 Q9DG82 BRARE	Q9DG82 brachydanio
23	1201	43.1	543	2 Q6NUJ1 HUMAN	Q6NUJ1 homo sapien
24	1153	41.3	240	2 Q5BJH1 HUMAN	Q5BJH1 homo sapien
25	1099.5	39.4	525	2 Q8C1C1 MOUSE	Q8C1C1 mus muscu
26	803	28.8	245	2 Q5Z162 CHICK	Q5Z162 gallus gall
27	786	28.2	449	2 Q8BUV5 MOUSE	Q8BUV5 mus muscu
28	772	27.7	402	2 Q8CLN0 MOUSE	Q8CLN0 mus muscu
29	558	20.0	253	2 Q5R848 POMPY	Q5R848 pongo pygma
30	552	19.8	953	2 Q9Y125 DROME	Q9Y125 drosophila
31	545	19.5	241	2 Q8N7T4 HUMAN	Q8N7T4 homo sapien

32	517	18.5	121	2	P79254 SHEEP	P79254 ovis aries
33	511	18.3	965	2	O15997 BOMMO	O15997 bombyx mori
34	497.5	17.8	478	2	Q7PJ67 ANOQA	Q7PJ67 anopheles g
35	479	17.2	922	2	Q5D981 SCHJA	Q5D981 schistosome
36	432.5	15.5	876	2	Q8IMH4 DROME	Q8IMH4 drosophila
37	407	14.6	533	2	Q4JHB2 TRIVA	Q4JHB2 trichomonas
38	400.5	14.4	890	2	Q7PMW6 ANOQA	Q7PMW6 anopheles g
39	393	14.1	80	1	SAP PIG	P81405 sus scrofa
40	382	13.7	377	1	PSPB_MOUSE	P50405 mus musculus
41	379	13.6	376	1	PSPB_RAT	P22355 rattus norv
42	378.5	13.6	613	2	Q4JM32 TRIVA	Q4JM32 trichomonas
43	377	13.5	376	2	Q6IN44 RAT	Q6IN44 rattus norv
44	349.5	12.5	378	2	O35489 CAVPO	O35489 cavia porce
45	349.5	12.5	381	1	PSPB_HUMAN	P07988 homo sapien

ALIGNMENTS

RESULT 1

ID	SAP_HUMAN	STANDARD:	PRT:	524 AA.
AC	P07602	P07292; P15793; P78538; P78541; P78546; P78547; P78558;		
AC	Q61B06	Q92739; Q92740; Q92741; Q92742;		
DT	01-APR-1988	(Rel. 07, Created)		
DT	13-SEP-2005	(Rel. 14, Last sequence update)		
DE	Proactivator polypeptide precursor (Contains: Saposin A (Protein A); Saposin B-Val; Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator); Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)).			
GN	Name=PSAP;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [MRNA].			
RC	TISSUE=Liver;			
RA	MEDLINE=90129043; PubMed=2515150;			
RT	Rorman E.G., Grabowski G.A.;			
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";			
RL	Genomics 5:486-492(1989).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [MRNA].			
RC	MEDLINE=89255151; PubMed=2498298;			
RA	Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;			
RT	"Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";			
RL	J. Biochem. 105:152-154(1989).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RA	Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;			
RT	"Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201).";			
RT	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SAP-MU-0).			
RC	TISSUE=Brain, Eye, and Skin;			
RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RT	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RT	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RT	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,			
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RT	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RT	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RT	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			

Wed Jan 18 11:58:41 2006

- RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
- [5] NUCLEOTIDE SEQUENCE OF 14-524.
- RP MEDLINE=88321660; PubMed=2842863;
- RA O'Brien J.S., Kretz K.A., Dewji N., Wenger D.A., Esch F., Fluharty A.L.; "Cloning of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic locus."; Science 241:1098-1101(1988).
- [6] NUCLEOTIDE SEQUENCE OF 14-524.
- RP MEDLINE=92307663; PubMed=1612590;
- RA Rorman E.G., Scheinker V., Grabowski G.A.; "Structure and evolution of the human prosaposin chromosomal gene."; Genomics 13:312-318(1992).
- [7] PROTEIN SEQUENCE OF 17-24; 165-172; 180-189 AND 298-302.
- RP MEDLINE=93311991; PubMed=8323276; DOI=10.1006/abbi.1993.1328;
- RA Hiraiwa M., O'Brien J.S., Kishimoto Y., Galdzicka M., Fluharty A.L., Ginns E.I., Martin B.M.; "Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposins (sphingolipid activator proteins)."; Arch. Biochem. Biophys. 304:110-116(1993).
- [8] PROTEIN SEQUENCE OF 17-26.
- RP TISSUE=Milk;
- RC MEDLINE=92068206; PubMed=1958198;
- RA Kondoh K., Hinenio T., Sano A., Kakimoto Y.; "Isolation and characterization of prosaposin from human milk."; Biochem. Biophys. Res. Commun. 181:286-292(1991).
- [9] NUCLEOTIDE SEQUENCE OF 59-125 AND 304-513.
- RP TISSUE=Brain;
- RC MEDLINE=91192146; PubMed=2013321; DOI=10.1016/0014-5793(91)80308-P;
- RA Holtzschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D., Suzuki K.; "The organization of the gene for the human cerebroside sulfate activator protein."; FEBS Lett. 280:267-270(1991).
- [10] PARTIAL PROTEIN SEQUENCE OF 60-142.
- RP MEDLINE=89240739; PubMed=2717620;
- RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S., Kishimoto Y.; "Saposin A: second cerebroside activator protein."; Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
- [11] PROTEIN SEQUENCE OF 62-84 AND 410-431.
- RP MEDLINE=93380576; PubMed=8370464; DOI=10.1016/0014-5793(93)80908-D;
- RA Tynnel J., Palmer D.N., Baumann M., Haltia M.; "Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis."; FEBS Lett. 330:8-12(1993).
- [12] NUCLEOTIDE SEQUENCE OF 164-524.
- RP MEDLINE=88068647; PubMed=2825202;
- RA Dewji N.N., Wenger D.A., O'Brien J.S.; "Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 precursor."; Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
- [13] NUCLEOTIDE SEQUENCE OF 195-263.
- RP MEDLINE=86130593; PubMed=2868718;
- RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F., Hill P., O'Brien J.S.; "Molecular cloning of the sphingolipid activator protein-1 (SAP-1), the sulfate sulfate activator."; Biochem. Biophys. Res. Commun. 134:989-994(1986).
- [14] PROTEIN SEQUENCE OF 195-274.
- RP MEDLINE=89207118; PubMed=3242555;
- RA Kleinschmidt T., Christomanou H., Braunitzer G.; "Complete amino-acid sequence of the naturally occurring A2 activator protein for enzymic sphingomyelin degradation: identity to the sulfate activator protein (SAP-1)."; Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
- [15] PROTEIN SEQUENCE OF 195-274.
- RP TISSUE=Kidney;
- RC MEDLINE=91006165; PubMed=2209618;
- RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.; "The complete amino-acid sequence of human ganglioside GM2 activator protein and cerebroside sulfate activator protein."; Eur. J. Biochem. 192:709-714(1990).
- [16] PROTEIN SEQUENCE OF 311-390.
- RP MEDLINE=88163077; PubMed=3442600;
- RA Kleinschmidt T., Christomanou H., Braunitzer G.; "Complete amino-acid sequence and carbohydrate content of the naturally occurring glucosylceramide activator protein (A1 activator) absent from a new human Gaucher disease variant."; Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
- [17] PARTIAL PROTEIN SEQUENCE OF 405-484.
- RP MEDLINE=89025876; PubMed=2845979;
- RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.; "Saposin D: a sphingomyelinase activator."; Biochem. Biophys. Res. Commun. 156:403-410(1988).
- [18] PROTEIN SEQUENCE OF 407-484.
- RP MEDLINE=89001190; PubMed=3048308;
- RA Furst W., Machleidt W., Sandhoff K.; "The precursor of sulfate activator protein is processed to three different proteins."; Biol. Chem. Hoppe-Seyler 369:317-328(1988).
- [19] PARTIAL PROTEIN SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
- RC TISSUE=Urine;
- RA MEDLINE=20032116; PubMed=10562467; DOI=10.1006/mgme.1999.2900;
- RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P., Waring A.J., To T., Fluharty C.B., Faull K.F.; "Preparation of the cerebroside sulfate activator (CSAct or saposin B) from human urine."; Mol. Genet. Metab. 68:391-403(1999).
- [20] DISULFIDE BONDS OF SAPOSINS B AND C, AND MASS SPECTROMETRY.
- RP MEDLINE=95247790; PubMed=7730378; DOI=10.1074/jbc.270.17.9953;
- RA Vaccaro A.M., Salvioli R., Barca A., Tatti M., Ciaffoni F., Maras B., Siciliano R., Zappacosta P., Amoresano A., Pucci P.; "Structural analysis of saposin C and B. Complete localization of disulfide bridges."; J. Biol. Chem. 270:9953-9960(1995).
- [21] STRUCTURE OF CARBOHYDRATE ON ASN-215.
- RP MEDLINE=21110404; PubMed=11180632;
- RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P., Stevens R.L., Fluharty C.B., Fluharty A.L.; "Structure of the asparagine-linked sugar chains of porcine kidney and human urine cerebroside sulfate activator protein."; J. Mass Spectrom. 35:1416-1424(2000).
- [22] SAPOSIN D DISULFIDE BONDS.
- RP MEDLINE=99337688; PubMed=10406958;
- RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,

RA Anoreano A., Vaccaro A.M.;
RT "Structural and membrane-binding properties of saposin D.";

Query Match 100.0%; Score 2789; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKS 60
DB 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKS 60

QY 61 LPDCDKDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPDCDKDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPFLLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPFLLY 180

QY 181 PDGPRSKPQKONGDVQDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADI 240
DB 181 PDGPRSKPQKONGDVQDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVI PALELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVI PALELVE 300

QY 301 PIKGEVPAKSDVYCEVCFVLEKVTGLIDNNKTEKEILDADFKMCKSLPKSISECQEV 360
DB 301 PIKGEVPAKSDVYCEVCFVLEKVTGLIDNNKTEKEILDADFKMCKSLPKSISECQEV 360

QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLGVYLD 420
DB 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLGVYLD 420

QY 421 RNLEKNSTQKIELAALKEGCSFLPDYPYQKQCFVAEYEPVLEILVEVMDPSFVCLKIG 480
DB 421 RNLEKNSTQKIELAALKEGCSFLPDYPYQKQCFVAEYEPVLEILVEVMDPSFVCLKIG 480

QY 481 ACPSAHKPLLGTEKICWGPSYWCNTETAACQNAVEHCKRHVYN 524
DB 481 ACPSAHKPLLGTEKICWGPSYWCNTETAACQNAVEHCKRHVYN 524

RESULT 2
Q5JQ36 HUMAN
ID Q5JQ36 HUMAN PRELIMINARY; PRT; 524 AA.
AC Q5JQ36;
DT 10-WAY-2005 (T-REMBLrel. 30, Created)
DT 10-WAY-2005 (T-REMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (T-REMBLrel. 31, Last annotation update)
DE Proboposin (Variant Gaucher disease and variant metachromatic leukodystrophy).
GN Name=PSAP; ORFNames=RP11-472K8.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBTaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor vector";
RT Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL731541; CAI40837.1; -; Genomic DNA.
DR EMBL; BT006849; AAP35495.1; -; mRNA.

DR EMBL; AC073370; CAI40837.1; JOINED; Genomic DNA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
SQ SEQUENCE 524 AA, 58112 MW, 71977P7A8C9E1533 CRC64;

Query Match 100.0%; Score 2789; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKS 60
DB 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKS 60

QY 61 LPDCDKDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPDCDKDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPFLLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPFLLY 180

QY 181 PDGPRSKPQKONGDVQDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADI 240
DB 181 PDGPRSKPQKONGDVQDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVI PALELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPMQTLVPAKVASKNVI PALELVE 300

QY 301 PIKGEVPAKSDVYCEVCFVLEKVTGLIDNNKTEKEILDADFKMCKSLPKSISECQEV 360
DB 301 PIKGEVPAKSDVYCEVCFVLEKVTGLIDNNKTEKEILDADFKMCKSLPKSISECQEV 360

QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLGVYLD 420
DB 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLGVYLD 420

QY 421 RNLEKNSTQKIELAALKEGCSFLPDYPYQKQCFVAEYEPVLEILVEVMDPSFVCLKIG 480
DB 421 RNLEKNSTQKIELAALKEGCSFLPDYPYQKQCFVAEYEPVLEILVEVMDPSFVCLKIG 480

QY 481 ACPSAHKPLLGTEKICWGPSYWCNTETAACQNAVEHCKRHVYN 524
DB 481 ACPSAHKPLLGTEKICWGPSYWCNTETAACQNAVEHCKRHVYN 524

RESULT 3
Q53FJ5 HUMAN
ID Q53FJ5 HUMAN PRELIMINARY; PRT; 524 AA.
AC Q53FJ5;
DT 13-SEP-2005 (T-REMBLrel. 31, Created)
DT 13-SEP-2005 (T-REMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-REMBLrel. 31, Last annotation update)
DE Proboposin (Variant Gaucher disease and variant metachromatic leukodystrophy) variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBTaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of eucaryotic mRNAs with oligoribonucleotides";
RL Gene 138:171-174 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane;
RA Suzuki Y., Yoshitomo K., Maruyama K., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-

Wed Jan 18 11:58:41 2006

```

end-enriched cDNA library." ;
Gene 200:149-156(1997).
[3]
NUCLEOTIDE SEQUENCE.
TISSUE=Synovial membrane;
Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
Tanaka A., Yokoyama S.;
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK232290; BAD97010.1; -, mRNA.
FT NON TER 1
SQ SEQUENCE 524 AA; 58140 MW; 7034F0C71C2226BC CRC64;

Query Match 99.8%; Score 2785; DB 2; Length 524;
Best Local Similarity 99.8%; Pred. No. 2.3e-173;
Matches 523; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVNKNPTVKS 60
Db 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVNKNPTVKS 60
QY 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLILY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLILY 180
QY 181 PQDGRSPKPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGGMADI 240
Db 181 PQDGRSPKPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKBEICALVGFCDVEKEMPMQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKBEICALVGFCDVEKEMPMQTLVPAKASKNVI PALELVE 300
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQEV 360
Db 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQEV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTOPKGGFCEVCKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTOPKGGFCEVCKLVGYLD 420
QY 421 RNLEKNSTQEIILAEKGCFLPDYQKQCFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTQEIILAEKGCFLPDYQKQCFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

RESULT 4
Q5JQ37 HUMAN PRELIMINARY; PRT; 559 AA.
ID Q5JQ37;
AC Q5JQ37;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Prosaposin (Variant Gaucher disease and variant metachromatic leukodystrophy).
GN Name=PSAP; ORFNames=RP11-472K9.1-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731541; CAI40836.1; -, Genomic_DNA.

end-enriched cDNA library." ;
Gene 200:149-156(1997).
[3]
NUCLEOTIDE SEQUENCE.
TISSUE=Synovial membrane;
Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
Tanaka A., Yokoyama S.;
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK232290; BAD97010.1; -, mRNA.
FT NON TER 1
SQ SEQUENCE 524 AA; 58140 MW; 7034F0C71C2226BC CRC64;

Query Match 99.8%; Score 2785; DB 2; Length 524;
Best Local Similarity 99.8%; Pred. No. 2.3e-173;
Matches 523; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVNKNPTVKS 60
Db 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVNKNPTVKS 60
QY 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLILY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLILY 180
QY 181 PQDGRSPKPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGGMADI 240
Db 181 PQDGRSPKPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKBEICALVGFCDVEKEMPMQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKBEICALVGFCDVEKEMPMQTLVPAKASKNVI PALELVE 300
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQEV 360
Db 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQEV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTOPKGGFCEVCKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTOPKGGFCEVCKLVGYLD 420
QY 421 RNLEKNSTQEIILAEKGCFLPDYQKQCFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTQEIILAEKGCFLPDYQKQCFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

RESULT 5
Q5JQ37 HUMAN PRELIMINARY; PRT; 530 AA.
ID Q5JQ37;
AC Q5JQ37;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Prosaposin variant (Fragment).
GN Name=prosaposin variant;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.

```



```

OX NCB1_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB209776; BAD93013.1; -, mRNA.
FT NON TER
SQ SEQUENCE 530 AA; 58727 MW; 6CA1F0159B182BC9 CRC64;

Query Match          99.6%; Score 2777.5; DB 2; Length 530;
Best Local Similarity 99.4%; Pred. No. 7.2e-173;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 4 MYALFLLASLLGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 63
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
Db 64 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 123
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNQKLESNKIPELDMTEVVAFPFMANIPILLY 180
Db 124 IKGEMSRPGEVCSALNLCESLQKHLAEHLNQKLESNKIPELDMTEVVAFPFMANIPILLY 183
Qy 181 PQDGPRSKPQPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGMWADI 240
Db 184 PQDGPRSKPQPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGMWADI 243
Qy 241 CKNYISOYSYATQMMHMH--OPKEICALVGFCDEVKEMPQMTLVPAKVASKNVIPALE 297
Db 244 CKNYISOYSYATQMMHMHQDQOPKEICALVGFCDEVKEMPQMTLVPAKVASKNVIPALE 303
Qy 298 LVPEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSEEC 357
Db 304 LVPEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSEEC 363
Qy 358 QEVVDYTGSSILSILLESVPSELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVG 417
Db 364 QEVVDYTGSSILSILLESVPSELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVG 423
Qy 418 YLDRLNLEKNSTKQEIILAALEKGCFLPDYPYQKCDQFVAEYEPVLIEILVEVMDPSFVCL 477
Db 424 YLDRLNLEKNSTKQEIILAALEKGCFLPDYPYQKCDQFVAEYEPVLIEILVEVMDPSFVCL 483
Qy 478 KIGACPSAHKPLLGTEKICWGSPSYWCNTETAACNAVEHCKRHVN 524
Db 484 KIGACPSAHKPLLGTEKICWGSPSYWCNTETAACNAVEHCKRHVN 530

RESULT 6
Q5NVD5_PONPY PRELIMINARY; PRT; 527 AA.
AC Q5NVD5;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFp459F0110.
GN Name=DKFp459F0110;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Pongo.
OX NCB1_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Pobo G., Han M., Wiemann S.;

```

```

RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR926102; CA129728.1; -, mRNA.
DR SMR; Q5NVD5; 194-275, 314-393.
DR GO: GO:0005764; C:lysosome; IEA.
DR GO: GO:0006629; P:lipid metabolism; IEA.
DR GO: GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapaB 1.
DR InterPro; IPR008138; SapaB 2.
DR InterPro; IPR008140; SapaB sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapaB 1; 4.
DR Pfam; PF03489; SapaB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapaB sub; 2.
DR SMART; SM00162; SAPA; 2.
KW Hypothetical protein.
SQ SEQUENCE 527 AA; 58469 MW; 293FBB746C29C4D0 CRC64;

Query Match          99.6%; Score 2776.5; DB 2; Length 527;
Best Local Similarity 99.2%; Pred. No. 8.3e-173;
Matches 523; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLLASLLGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNQKLESNKIPELDMTEVVAFPFMANIPILLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAEHLNQKLESNKIPELDMTEVVAFPFMANIPILLY 180
Qy 181 PQDGPRSKPQPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGMWADI 240
Db 181 PQDGPRSKPQPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGMWADI 240
Qy 241 CKNYISOYSYATQMMHMH--OPKEICALVGFCDEVKEMPQMTLVPAKVASKNVIPALE 297
Db 241 CKNYISOYSYATQMMHMHQDQOPKEICALVGFCDEVKEMPQMTLVPAKVASKNVIPALE 300
Qy 298 LVPEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSEEC 357
Db 301 LVPEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSEEC 360
Qy 358 QEVVDYTGSSILSILLESVPSELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVG 417
Db 361 QEVVDYTGSSILSILLESVPSELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVG 420
Qy 418 YLDRLNLEKNSTKQEIILAALEKGCFLPDYPYQKCDQFVAEYEPVLIEILVEVMDPSFVCL 477
Db 421 YLDRLNLEKNSTKQEIILAALEKGCFLPDYPYQKCDQFVAEYEPVLIEILVEVMDPSFVCL 480
Qy 478 KIGACPSAHKPLLGTEKICWGSPSYWCNTETAACNAVEHCKRHVN 524
Db 481 KIGACPSAHKPLLGTEKICWGSPSYWCNTETAACNAVEHCKRHVN 527

RESULT 7
Q5R4U7_PONPY PRELIMINARY; PRT; 526 AA.
AC Q5R4U7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFp459J1314.
GN Name=DKFp459J1314;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Bloeker H., Boecker H., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Oeanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR861144; CAH93219.1; -; mRNA.
DR SMR; Q8407; 194-274, 313-392.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapaB.1.
DR InterPro; IPR008138; SapaB.2.
DR InterPro; IPR008140; SapaB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapaB.1; 4.
DR Pfam; PF03489; SapaB.2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapaB_sub; 2.
DR SMART; SM00162; Sapa; 2.
KW Hypothetical protein.
SQ SEQUENCE 526 AA; 58325 MW; CF3B146DDB6F5539 CRC64;

Query Match 99.3%; Score 2770; DB 2; Length 526;
Best Local Similarity 99.2%; Pred. No. 2.2e-172;
Matches 522; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 MYALFLASLLGALAGPVLGKCTGRSAVWCNQVKTASDCGAVKHCLQTVNNKPTVKS 60
DB 1 MYALFLASLLGALAGPVLGKCTGRSAVWCNQVKTASDCGAVKHCLQTVNNKPTVKS 60

QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNNMSASCKEIVDSYLPVLDI 120
DB 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNNMSASCKEIVDSYLPVLDI 120

QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLY 180

QY 181 PDGPRSKPQPKDNGDVQCDCIQMTDIOQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 181 PDGPRSKPQPKDNGDVQCDCIQMTDIOQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMH--OPKEICALVGFCDEVKEMPQTLVPAKVASKNVIPALEL 298
DB 241 CKNYISQYSEIAIQMMHMHDDQPKKEICALVGFCDEVKEMPQTLVPAKVASKNVIPALEL 300

QY 299 VEPFKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQ 358
DB 301 VEPFKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQ 360

QY 359 EVVDTYGSILSLLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLGVY 418
DB 361 EVVDTYGSILSLLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLGVY 420

QY 419 LDRNLEKNSTKQBIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPVCLK 478
DB 421 LDRNLEKNSTKQBIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPVCLK 480

QY 479 IGACPSAHKPLLTGTEKICNGSPSYWCQNTETAACNAVEHCCKRHVN 524
DB 481 IGACPSAHKPLLTGTEKICNGSPSYWCQNTETAACNAVEHCCKRHVN 526

RESULT 8
Q4R590_MACFA

ID Q4R590_MACFA PRELIMINARY; PRT; 497 AA.
AC Q4R590;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Brain cDNA, clone: QccE-13090, similar to human prosaposin (variant
DE Gaucher disease and variantmetachromatic leukodystrophy) (PSAP),
DE (Brain cDNA, clone: QccE-13989, similar to human prosaposin (variant
DE Gaucher disease and variantmetachromatic leukodystrophy)
DE (PSAP)).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hitata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution;
RT Comparative analysis between human and cynomolgus monkey cDNAs";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB169654; BAB01735.1; -; mRNA.
DR EMBL; AB169527; BAE01609.1; -; mRNA.
SQ SEQUENCE 497 AA; 55306 MW; 4A1974F8DB883900 CRC64;

Query Match 87.9%; Score 2451.5; DB 2; Length 497;
Best Local Similarity 97.9%; Pred. No. 1.3e-151;
Matches 468; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 1 MYALFLASLLGALAGPVLGKCTGRSAVWCNQVKTASDCGAVKHCLQTVNNKPTVKS 60
DB 1 MYALFLASLLGALAGPVLGKCTGRSAVWCNQVKTASDCGAVKHCLQTVNNKPTVKS 60

QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNNMSASCKEIVDSYLPVLDI 120
DB 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNNMSASCKEIVDSYLPVLDI 120

QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLY 180

QY 181 PDGPRSKPQPKDNGDVQCDCIQMTDIOQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 181 PDGPRSKPQPKDNGDVQCDCIQMTDIOQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMH--OPKEICALVGFCDEVKEMPQTLVPAKVASKNVIPALE 297
DB 241 CKNYISQYSEIAIQMMHMHDDQPKKEICALVGFCDEVKEMPQTLVPAKVASKNVIPALE 300

QY 298 VEPFKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEEC 357
DB 301 VEPFKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEEC 360

QY 358 QEVVDTYGSILSLLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLGVY 417
DB 361 QEVVDTYGSILSLLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLGVY 420

QY 418 YLDRNLEKNSTKQBIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPFV 475
DB 421 YLDRNLEKNSTKQBIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPFV 478

RESULT 9
SAP_BOVIN
ID - SAP_BOVIN STANDARD; PRT; 525 AA.
AC P26779; Q9N2G4;

Db	241	MCKNYINQYSEVAIQVMHMQPKEICVLAFGFCDEVKEMPMTLVPAAEVVSENVTIPALGLV	300
Qy	300	EPKKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSEECQE	359
Db	301	EPKKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSEECQE	360
Qy	360	VVDYTGSSILILAEVSELPVCSMLHLCSGTRLPALTTHVHTQPKDGGFCVCKKLVGYL	419
Db	361	VVDYTGSSILILAEVSELPVCSMLHLCSGTRLPALTTHVHTQPKDGGFCVCKKLVGYL	420
Qy	420	DRNLEKSTKQEILAALEKSGSFLPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLIKI	479
Db	421	DRNLEKSTKQEILAALEKSGSFLPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLIKI	480
Qy	480	GACPSAHKPLGTEKICNGPSPYWCQNTTAAQCNVHECHKRVWN	524
Db	481	GACPSAHKPLGTEKICNGPSPYWCQNTTAAQCNVHECHKRVWN	525
RESULT 10			
Qy	QSR406	PONPY PRELIMINARY; PRT; 452 AA.	
AC	QSR406		
DT	01-FEB-2005	(TremBLrel. 29, Created)	
DT	01-FEB-2005	(TremBLrel. 29, Last sequence update)	
DT	01-FEB-2005	(TremBLrel. 29, Last annotation update)	
DE	Hypothetical protein DKFZp459G152.		
GN	Name=DKFZp459G152;		
OS	Pongo pygmaeus (Orangutan).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Pongo.		
OX	NCBI_TaxID=9600;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Cortex;		
RG	The German cDNA Consortium;		
RA	Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,		
RA	Fobo G., Han M., Wiemann S.;		
RL	Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; CR861454; CAH93510.1; -, mRNA.		
DR	SMR; QSR406; 119-200, 239-318.		
DR	GO; GO:0005764; C:lysosome; IEA.		
DR	GO; GO:0006629; P:lipid metabolism; IEA.		
DR	GO; GO:0006665; P:sphingolipid metabolism; IEA.		
DR	InterPro; IPR003119; Sapa.		
DR	InterPro; IPR007856; Sapa 1.		
DR	InterPro; IPR008138; Sapa 2.		
DR	InterPro; IPR008140; Sapa sub.		
DR	InterPro; IPR008373; Saposin.		
DR	InterPro; IPR008139; SaposinB.		
DR	Pfam; PF02199; Sapa; 1.		
DR	Pfam; PF05184; Sapa 1; 4.		
DR	Pfam; PF03489; Sapa 2; 4.		
DR	PRINTS; PR01797; SAPOSIN.		
DR	ProDom; PD001732; Sapa sub; 2.		
DR	SMART; SM00162; Sapa; 1.		
KW	Hypothetical protein.		
SQ	SEQUENCE 452 AA; 50664 MW; 6D30EAC665A399DD CRC64;		
Query Match 85.1%; Score 2374.5; DB 2; Length 452;			
Best Local Similarity 99.1%; Pred. No. 1.2e-146;			
Matches 448; Conservative 1; Mismatches 0; Indels 3; Gaps 1;			
Qy	76	MLKDNATEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKNGSRPGEVCSAL	135
Db	1	MLKDNATEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKNGSRPGEVCSAL	60
Qy	136	NLCESLQKHLAEHLNHOLESNKIPDLMTVEVAPFWANIPLLYPODGRSKPQKNG	195
Db	61	NLCESLQKHLAEHLNHOLESNKIPDLMTVEVAPFWANIPLLYPODGRSKPQKNG	120
Qy	196	DVCDQCIMQVTDIQTAVRTNSTFQVALVEHVEKCDRLGPGMADICKNYISQYSEIAIQM	255

Db	121	DVCDQCIMQVTDIQTAVRTNSTFQVALVEHVEKCDRLGPGMADICKNYISQYSEIAIQM	180
Qy	256	MMHM--QPKKEICALVGFCDVEKEMPMTLVPAAKVNIPALVELPEPIKKEHVPKASD	312
Db	181	MMHMDDQPKKEICALVGFCDVEKEMPMTLVPAAKVNIPALVELPEPIKKEHVPKASD	240
Qy	313	VYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSEECQEVVDYTGSSILSIL	372
Db	241	VYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSEECQEVVDYTGSSILSIL	300
Qy	373	LEEVSPVLVCSMLHLCSGTRLPALTTHVHTQPKDGGFCVCKKLVGYLDRLNLEKSTKQEI	432
Db	301	LEEVSPVLVCSMLHLCSGTRLPALTTHVHTQPKDGGFCVCKKLVGYLDRLNLEKSTKQEI	360
Qy	433	LALEKSGSFLPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLIKIGACPSAHKPLGLT	492
Db	361	LALEKSGSFLPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLIKIGACPSAHKPLGLT	420
Qy	493	EKCIWGPSYWCQNTTAAQCNVHECHKRVWN	524
Db	421	EKCIWGPSYWCQNTTAAQCNVHECHKRVWN	452
RESULT 11			
SAP RAT			
ID - SAP RAT	STANDARD; PRT; 554 AA.		
AC	P10560; Q62841; Q64190;		
DT	01-JUL-1989 (Rel. 11, Created)		
DT	01-JUL-1989 (Rel. 11, Last sequence update)		
DT	13-SEP-2005 (Rel. 46, Last annotation update)		
DE	Sulfated glycoprotein 1 precursor (SGP-1) (Proasaposin).		
GN	Name=Psap; Synonyms=SGp1;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Murioidea; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.		
RC	TISSUE=Sertoli cell;		
RX	MEDLINE=89000647; PubMed=3048385;		
RA	Collard M.W., Sylvester S.R., Tsuruta J.K., Griewald M.D.;		
RT	"Bioynthesis and molecular cloning of sulfated glycoprotein 1		
RT	secreted by rat Sertoli cells: sequence similarity with the 70-		
RT	kilodalton precursor to sulfatide/GM1 activator.";		
RL	Biochemistry 27:4557-4564(1988).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [MRNA].		
RC	TISSUE=Testis;		
RX	MEDLINE=96128541; PubMed=8573994;		
RA	Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;		
RT	"Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in		
RT	Sertoli cells.";		
RL	Histol. Histopathol. 10:1023-1034 (1995).		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE [MRNA].		
RC	TISSUE=Testis;		
RX	MEDLINE=96175245; PubMed=8601692;		
RA	Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;		
RT	"Expression and tissue distribution of rat sulfated glycoprotein-1		
RT	(prosaposin).";		
RL	J. Histochem. Cytochem. 44:327-337(1996).		
CC	-1- SUBCELLULAR LOCATION: Extracellular.		
CC	-1- SIMILARITY: Contains 2 saposin A-type domains.		
CC	-1- SIMILARITY: Contains 4 saposin B-type domains.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		

DR EMBL; M19936; AAA42136.1; -; mRNA.
DR EMBL; S81353; AAB36042.2; -; mRNA.
DR EMBL; S81373; AAB36223.2; -; mRNA.
DR PIR; A28716; A28716.
DR KSSP; Q22739; I1N65.
DR SMR; P10960; 195-271.
DR Ensemble; ENSRNOG00000000571; Rattus norvegicus.
DR RGD; 3423; Peap.
DR GO; GO:0046836; P-glycolipid transport; NAS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB 1.
DR InterPro; IPR008138; SapB 2.
DR InterPro; IPR008140; SapB sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapB 1; 3.
DR Pfam; PF03489; SapB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB sub; 1.
DR PROSITE; PS11110; SAP_A; 2.
DR PROSITE; PS50015; SAP_B; 4.
KW Direct protein sequencing; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 554 Sulfated glycoprotein 1.
FT DOMAIN 18 58 Saposin A-type 1.
FT DOMAIN 59 142 Saposin B-type 1.
FT DOMAIN 193 274 Saposin B-type 3.
FT DOMAIN 310 391 Saposin B-type 3.
FT DOMAIN 435 516 Saposin B-type 4.
FT DOMAIN 518 554 Saposin A-type 2.
FT CARBOHYD 80 80 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 214 214 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 456 456 N-linked (GlcNAc...) (Potential).
FT DISULFID 63 138 By similarity.
FT DISULFID 66 132 By similarity.
FT DISULFID 94 106 By similarity.
FT DISULFID 197 270 By similarity.
FT DISULFID 200 264 By similarity.
FT DISULFID 229 240 By similarity.
FT DISULFID 314 387 By similarity.
FT DISULFID 317 381 By similarity.
FT DISULFID 345 356 By similarity.
FT DISULFID 439 512 By similarity.
FT DISULFID 442 506 By similarity.
FT DISULFID 470 481 By similarity.
FT CONFLICT 115 115 p -> L (in Ref. 2 and 3).
FT CONFLICT 299 299 D -> E (in Ref. 2).
FT CONFLICT 462 462 I -> V (in Ref. 3).
FT CONFLICT 527 527 W -> R (in Ref. 3).
FT CONFLICT 536 536 S -> M (in Ref. 3).
SQ SEQUENCE 554 AA; 61124 MW; DFE3F3A3A0520C6B CRC64;

Query Match
Best Local Similarity 71.68; Score 1996; DB 1; Length 554;
Matches 369; Conservative 66.38; Pred. No. 7.6e-122; Mismatches 78; Indels 32; Gaps 3;

QY 1 MYALFLIASLIGALAGVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTAVNKPVTKS 60
DB 1 MYALALLASLLVLTALTSFVQDPKICSGSAVCRVDKTAIDCRVHKCQQWWSKPTAKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATREIILVLEKTCWLPKPNMGSACKEIVDSYLPVLDI 120
DB 61 LPDCICKTVTEAGNLKDNATREIILVLEKTCWTHDSLSASCKEVVDSYLPVLDM 120
QY 121 IKGEMSPGVCNALNICSLOKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180
DB 121 IKGEMSPGVCNALNICSLOQLQYLAEQN-QRQLESNKIPFVDLAVVVPFMSNIPILLY 179
QY 181 PQDPRSPQPPKONGDVQCDCIOMVTDIQTAVRTNSTFVQALVEHVKEECDRLPGGMADI 240

Db 180 PQDPRSPQPPKANEDVQCDCMKLVTDIQTAVRTNSTSFVQGLVDHVKEDCDRLPGGVSDI 239
QY 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDCEVEMPMQTLVPAKVASKNVIPALELVE 300
Db 240 CKNYDYDYSEVAVQMMHMQPKEICVMVGFCDCEVKEVPMRTLVPAEAIKNILPALELTD 299
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDADFMCKSKLPKSLSEECQEV 360
Db 300 PVEQDVIQAQWIFCQVCLVMKLSLSELIINNATELLINKLSKACSLLPAPASTKCOEV 359
QY 361 VDTYSSILSILLREVSPELVCSMLHLCSSG-----TRLP----- 394
Db 360 LVTFGSLDLVLMHEVNPFLGCVISCSANPMLVGLTPQAAAIVSALPKPEAPPKQPE 419
QY 395 -----ALTVHTVTPQKDGFGCEVCKKLVGLYLDRLNLEKNSTKQIEILAALEKSCSLPDPYQK 449
Db 420 EPKQSALRAHVPQKNGGFCVCKKLVILEHNLEKNSTKEIEILAALEKSCSLPDPYQK 479
QY 450 QCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAHKPLLGTEKINGSPSYWCONTETA 509
Db 480 QCDEFVARYEPLILLEILVEVMDPSFVCSKIGVCPSPAYKLLLGTEKCVMGFGYWCNSETA 539
QY 510 ACQNAVEHCKRHVMN 524
Db 540 ARCNVADHCKRHVMN 554

RESULT 12
Q6P7A4 RAT PRELIMINARY; PRT; 553 AA.
AC Q6P7A4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Prosaposin.
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh P.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gay L.J., Hulyk S.W.,
Richards D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Villalon D., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061759; AAH61759.1; -; mRNA.
DR SMR; Q6P7A4; 195-271.
DR GO; GO:0005764; C:lysosome; IEA.

DR	GO; GO:0006299; P:lipid metabolism; IEA.
DR	GO; GO:0006655; P:sphingolipid metabolism; IEA.
DR	InterPro; IPR003119; SApA.
DR	InterPro; IPR007856; SapB_1.
DR	InterPro; IPR008138; SapB_2.
DR	InterPro; IPR008140; SapB_sub.
DR	InterPro; IPR008373; Saposin.
DR	InterPro; IPR008139; SaposinB.
DR	pFam; PF05184; SapB_1; 3.
DR	pFam; PF05184; SapB_2; 4.
DR	pFam; PF03489; SapB_2; 4.
DR	PRINTS; PR01797; SAPOSIN.
DR	ProdDom; PD001732; SapB_sub; 1.
DR	SMART; SM00162; SAPA; 2.
SQ	SEQUENCE 553 AA; 61039 MW; 63F3DD5E0C523393 CRC64;
	Query Match 71.1%; Score 1982.5; DB 2; Length 553;
	Best Local Similarity 66.3%; Pred. No. 5.7e-121;
	Matches 368; Conservative 75; Mismatches 79; Indels 33; Gaps 4;
QY	1 MYALFLASLLGAALAGPVLGLKCEKTGSAVWCNQVKVTASDCGAVRGCLQTVMNKPVTKS 60
DB	1 MYALALLASLLVTALTISPVDPDKLSGGSAAVCVDKTAVIDCRVAKHCQQMWSKPTAKS 60
QY	61 LPDICKDVVAAGMLKDNAMEBEILLVLEKTCOMLKPKNMSASCKEIVDSYLPVILDI 120
DB	61 LPDICKTVTVEAGNLKKDNATEBEILLHYLEKTCWIIHSDSLSSASCKEVDSYLPVILDM 120
QY	121 IKGEMSPGEVCSALNLCESLOKLAEHLNHQLESNKIPELDTEVVAPPMANIPILLY 180
DB	121 IKGEMSPGEVCSALNCQLSEYLAQN-QRQLESNKIPEVDLARVVAFPMNSIPLLY 179
QY	181 PQDGPRSKPQKNDGVQCDCIQMVTDIQTAVRTNSTFVALVHVKEECDRLGPGMADI 240
DB	180 PQDPRSQPKANEDVCQCKMLVDIQTAVRTNSSVFQGLVHVKECDRLGFGVSDI 239
QY	241 CKNYISOYSETAIQMHHMPKEICALVGFCDEVKEMPQTLVPKAVASKNVIPALELVE 300
DB	240 CKNYVDQYSEVAVQMMHMPPKEICVMVGFCDVKRVPMRTLVPATEAIKNILPALELT 299
QY	301 PIKHGEVPAKSDVCEVCEPFVKEVTKLIDNNKTEKELDAFDKWCCKLPKLSSECQEV 360
DB	300 PY-EDVIQAQNVIFQCVQLVMRKSELIIINNATEELLIKGLSKACSLPALPASTKCQEV 358
QY	361 VDTYGSSITSLILEVSPDELVCSSLHLCSG-----TRLP----- 394
DB	359 LVTFGPSLDLVMEHVNFNFLCGVISLSANPNLVGTLEEQPAAIVTSALPREAPPKQPE 418
QY	395 -----ALTVHTVTPQKGGFCFEVKGLVGYLDRLNEKNTKQEIILAALKSGSFDPDPYOK 449
DB	419 EPKQSALRAHVPPQKNGGFCFEVKGLKVITYLEHNLEKNTKSTKEIILAALKSGSFDPDPYOK 478
QY	450 QCDQFVARYEPVLBIIVEMDPSFVCLKIGACPSAHKPLTGTEKICIGSPSVWCONTETA 509
DB	479 QCDEFVARYEPVLBIIVEMDPSFVCSKIGVCPYSAYKLLLTGTEKICVWGPGYWCQNMTA 538
QY	510 AQCNAVECHKRHVN 524
DB	539 ARCNADVCHKRHVN 553
RESULT 13	
QB8F01_MOUSE	PRELIMINARY; PRT; 554 AA.
ID	QB8F01_MOUSE PRELIMINARY; PRT; 554 AA.
AC	QB8F01; AC
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE	Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE	length enriched library, clone:E430014B12 product:proasoposin, full-
DE	insert sequence (Mus musculus kidney CCL-142 RAG cDNA, RIKEN full-
DE	length enriched library, clone:G430050G07 product:proasoposin, full
DE	insert sequence).


```

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NOD, and BALB/C; TISSUE=Kidney, and Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NOD, and BALB/C; TISSUE=Kidney, and Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayaashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK088369; BAC40308.1; -; mRNA.
DR ENBL; AK089998; BAC41035.1; -; mRNA.
DR HSSP; Q92739; IN69.
DR SMR; Q8BFQ1; 195-271.
DR MGI; MGI:197783; Pesp.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005739; C:mitchondrion; IDA.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB.1.
DR InterPro; IPR008138; SapB.2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SapA; 2.
DR Pfam; PF05184; SapB.1; 4.
DR Pfam; PF03489; SapB.2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00741; SapB; 4.
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;

Query Match 69.9%; Score 1950; DB 2; Length 554;
Best Local Similarity 64.0%; Pred. No. 7.6e-119;
Matches 355; Conservative 78; Mismatches 90; Indels 32; Gaps 2;

QY 1 MYALFLASLLGALAGPVLGLKCTRGSAVMQNVKTASDCGAVKHCLQTVNKKPTVKS 60
DB 1 MYALALFASLLATALTSPVDPTKSGGSAVLCDVKTAIDCGAVKHQQMWSKPTAKS 60

QY 61 LPCDICVVTAGMDKDNATEEELVLYLEKTCWMLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPCDICVVTAGNLLKNATQEEILVLYLEKTCWMLHDSLSASCKEVVDSYLPVILDM 120

QY 121 IKGEMSRPEVCNALNCSLSQKHLAELNHQKLESNKIPELDMTEVVAFFWANIPLLIY 180
DB 121 IKGEMSNPGEVCNALNCSLSQKLEQYLAQN-QKLESNKIPEVDMARVAPFMSNIPLLIY 179

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC	EMBL	S36200	AB22175.1	;	;	mRNA.
CC	EMBL	S71616	AB331059.1	; <td>;<td>mRNA.</td></td>	; <td>mRNA.</td>	mRNA.
DR	EMBL	U27340	AA92567.1	; <td>;<td>mRNA.</td></td>	; <td>mRNA.</td>	mRNA.
DR	EMBL	U57999	AA802695.1	; <td>;<td>Genomic_DNA.</td></td>	; <td>Genomic_DNA.</td>	Genomic_DNA.
DR	PIR	JH0604	JH0604			
DR	HSSP	Q92739	IN69			
DR	SMR	Q61207	195-274			
DR	Ensembl	ENSMUSG00000004207	Mus musculus.			
DR	MGI	MGI:97782	Esap			
DR	GO	GO:0005615	C:extracellular space			TAS.
DR	GO	GO:0005739	C:mitochondrion			IDA.
DR	InterPro	IPR009007	Pept_Aspartc_cat.			
DR	InterPro	IPR003119	SapA.			
DR	InterPro	IPR007856	SapB.1			
DR	InterPro	IPR008138	SapB.2			
DR	InterPro	IPR008140	SapB_sub.			
DR	InterPro	IPR008373	SaposinB.			
DR	InterPro	IPR008139	SaposinB.			
DR	Pfam	PF02199	SapB.1			4.
DR	Pfam	PF05184	SapB.2			4.
DR	Pfam	PF03489	SapB.2			4.
DR	PRINTS	PR01797	SAFOSIN.			
DR	ProDom	PD001732	SapB_sub.1			
DR	PROSITE	PS51110	SAP A; 2.			
DR	PROSITE	PS50015	SAP B; 4.			
DR	GlycoSite	Repeat	Signal.			
KW	Glycoprotein	1	16			
FT	SIGNAL	1	16			
FT	CHAIN	17	557			
FT	DOMAIN	18	58			
FT	DOMAIN	59	142			
FT	DOMAIN	193	277			
FT	DOMAIN	313	394			
FT	DOMAIN	438	519			
FT	DOMAIN	521	557			
FT	CARBOHYD	80	80			
FT	CARBOHYD	214	214			
FT	CARBOHYD	334	334			
FT	CARBOHYD	459	459			
FT	DISULFID	63	138			
FT	DISULFID	66	132			
FT	DISULFID	94	106			
FT	DISULFID	197	273			
FT	DISULFID	200	267			
FT	DISULFID	229	240			
FT	DISULFID	317	390			
FT	DISULFID	320	384			
FT	DISULFID	348	359			
FT	DISULFID	442	515			
FT	DISULFID	445	509			
FT	DISULFID	473	484			
FT	CONFLICT	83	83			
FT	CONFLICT	158	158			
FT	CONFLICT	160	160			
FT	CONFLICT	171	172			
FT	CONFLICT	244	244			
FT	CONFLICT	254	254			
FT	CONFLICT	255	255			
FT	CONFLICT	260	262			
FT	CONFLICT	307	307			
FT	CONFLICT	322	322			
FT	CONFLICT	349	350			
FT	CONFLICT	367	367			
FT	CONFLICT	370	370			
FT	CONFLICT	373	373			
FT	CONFLICT	391	391			
FT	CONFLICT	393	393			
FT	CONFLICT	406	406			
FT	CONFLICT	430	430			
FT	CONFLICT	430	430			

[illegible]

RESULT 15	SAP CHICK	STANDARD;	PRT;	518 AA.
ID	- SAP CHICK			
AC	013035;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;			
DE	Saposin C; Saposin D].			
GN	Name=PSAP;			
OS	Gallus gallus (Chicken)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
[1]	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 194-203.			
RP	TISSUE=Brain, and Liver;			
RC	MEDLINE=98129745; PubMed=9461526;			
RX	Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,			
RA	Faton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.;			
RA	"Cloning, expression and map assignment of chicken prosaposin.";			
RT	Biochem. J. 330:321-327(1998).			
RL	[2]			

RP NUCLEOTIDE SEQUENCE.
RA Altman N., Horowitz M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate (By similarity).
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC -!- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB003471; BAA19914.1; -; mRNA.
DR EMBL; AF108656; AAF05899.1; -; mRNA.
DR HSSP; Q92739; IN69.
DR Ensembl; ENSGALG00000004769; Gallus gallus.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapoB_sub; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00741; Saps; 4.
DR PROSITE; PS51110; SAP A; 2.
DR PROSITE; PS50015; SAP B; 4.
DR Direct protein sequencing; Glycoprotein; GM2-gangliosidosis;
DR Lipid metabolism; Lysosome; Repeat; Signal; Sphingolipid metabolism.
FT SIGNAL 1 17 Potential.
FT PROPEP 18 60
FT CHAIN 61 143 Saposin A.
FT PROPEP 145 193
FT CHAIN 194 276 Saposin B.
FT PROPEP 278 305
FT CHAIN 307 387 Saposin C.
FT PROPEP 389 398
FT CHAIN 399 480 Saposin D.
FT PROPEP 482 518
FT DOMAIN 19 59 Saposin A-type 1.
FT DOMAIN 61 143 Saposin B-type 1.
FT DOMAIN 194 277 Saposin B-type 2.
FT DOMAIN 307 388 Saposin B-type 3.
FT DOMAIN 399 480 Saposin B-type 4.
FT DOMAIN 482 518 Saposin A-type 2.

FT CARBOHYD	81	N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD	214	N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD	328	N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD	420	N-linked (GlcNAc. . .) (Potential).
FT DISULFID	64	By similarity.
FT DISULFID	67	By similarity.
FT DISULFID	95	By similarity.
FT DISULFID	197	By similarity.
FT DISULFID	200	By similarity.
FT DISULFID	229	By similarity.
FT DISULFID	311	By similarity.
FT DISULFID	314	By similarity.
FT DISULFID	342	By similarity.
FT DISULFID	403	By similarity.
FT DISULFID	406	By similarity.
FT DISULFID	434	By similarity.
FT CONFLICT	94	R -> T (in Ref. 2).
FT CONFLICT	486	E -> D (in Ref. 2).
SQ SEQUENCE	518 AA; 57601 MW; B803000E991C3963 CRC64;	

Query Match 60.2%; Score 1679.5; DB 1; Length 518;
Best Local Similarity 57.4%; Pred. No. 3.1e-101;
Matches 301; Conservative 96; Mismatches 114; Indels 13; Gaps 5;

Qy	4	LFLASLILGAALAGPVLGLKECTRGSAVWCQNVKTSACGAVKHCLQTVWKNKPTVKSLPC	63
Db	5	LLTLGLLAARVASPVLWQKCAKGPVWCQSLRTASOCGAVKHCCQNVWSPAVNSIPC	64
Qy	64	DICKVVTAAGDMLKDNATEEEILVLEKTCDWLPKNMSCKEIVDSYLPVILDIKNG	123
Db	65	DLCKELTVVGVKLVKNGTDEIRSYLEKRCFELPDQGLASECKEIVDSYLPVINDMIKE	124
Qy	124	EMSPRGVCSALNLCESLQKLALNKHQKQLESNKIPELDMTEVVAVPMANIPLLLYPOD	183
Db	125	EPDKPEVVCSSALSQSLQKHLAAMKQLQSNKIPELDFSELTSPFMANVPLLYPOD	184
Qy	184	GPRSKPQPKONGDYQCDCIQMWTDIQTAVRTNSTFVQALVHVKEECDRLGGMADICKN	243
Db	185	KPKQK--SKATEDVQCDCIRLVTVQEAVRTNATFVKSVAHAKEECDRLGGMDSMCKS	242
Qy	244	YISQYSEIAIQMMHM---QPKICALVGFCDVEKEMPQTLVPAKASKNVIPALELVE	300
Db	243	YISEYDLAIQMMHMKDQPKDICAMVGFCSVKSVPQLTLVPAQVVHE-----VKME	296
Qy	301	PIKHEVPAKSDVYCEVCEFLVKVTKLIDNNKTEKEILDAPDKMCKSLPKSLSEECQEV	360
Db	297	TVEKATVQEKTFVSVCEITMVKVEVTGLLESNKTEEEIVHEMEVVCYLLPASVKDQCKDF	356
Qy	361	VDYGSLSILILEVSPVSLVCSMLHLCSTGLPALTVHTVTPKDGKGFCECKLVGYLD	420
Db	357	IEVYQALIDMLLEATNPEAVCVMLKCAANKPQQPV--VKPA--GGFCDICRMIVAYAD	414
Qy	421	RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEMDPSFVCLKIG	480
Db	415	KELEKNATTTEIEALLEKVKCHFLPESVDQCVQVEQYEPVYVQVLLAEMMDPTFVCTKLG	474
Qy	481	ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN	524
Db	475	VCGAACKPLLGEDACVWNGPGYCKNMETAAQCNADVDCRRHVN	518

Search completed: January 13, 2006, 16:34:21
Job time : 202.536 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:26:00 ; Search time 54.6556 Seconds
(without alignments)
792.637 Million cell updates/sec

Title: US-10-801-517-1
Perfect score: 2789
Sequence: 1 MYALFLASLILGALAGPVL.....NTETAAQNAVEHCKRHVWN 524

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pdp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pdp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pdp:*
4: /cgn2_6/ptodata/1/iaa/PCTRUS COMB.pdp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pdp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	2 US-09-352-548-1	Sequence 1, Appl1
2	2789	100.0	524	2 US-09-949-016-6272	Sequence 6272, Ap
3	2789	100.0	524	2 US-08-928-074-23	Sequence 23, Appl
4	2789	100.0	535	2 US-09-949-016-8603	Sequence 8603, Ap
5	2766.5	99.2	523	1 US-08-100-247-2	Sequence 2, Appl1
6	2766.5	99.2	523	2 US-08-756-031-2	Sequence 2, Appl1
7	2759.5	98.9	523	1 US-08-232-513A-3	Sequence 3, Appl1
8	2757.5	98.9	523	1 US-08-483-146A-2	Sequence 2, Appl1
9	2757.5	98.9	523	1 US-08-484-594A-2	Sequence 2, Appl1
10	2757.5	98.9	523	2 US-08-076-258A-2	Sequence 2, Appl1
11	431	15.5	81	2 US-09-352-548-2	Sequence 2, Appl1
12	418.5	15.0	80	1 US-08-584-671-15	Sequence 15, Appl
13	418.5	15.0	80	2 US-09-027-376-15	Sequence 15, Appl
14	415.5	14.9	80	2 US-09-094-192-15	Sequence 15, Appl
15	412	14.8	80	1 US-08-100-247-3	Sequence 3, Appl1
16	412	14.8	80	1 US-08-483-146A-3	Sequence 3, Appl1
17	412	14.8	80	1 US-08-232-513A-4	Sequence 4, Appl1
18	412	14.8	80	1 US-08-484-594A-3	Sequence 3, Appl1
19	412	14.8	80	2 US-08-076-258A-3	Sequence 3, Appl1
20	412	14.8	80	2 US-08-756-031-3	Sequence 3, Appl1
21	412	14.8	80	2 US-08-928-074-24	Sequence 24, Appl
22	350.5	12.6	381	1 US-08-193-877-2	Sequence 2, Appl1
23	349.5	12.5	381	2 US-09-949-016-10057	Sequence 10057, A
24	327	11.7	61	1 US-08-584-671-13	Sequence 13, Appl
25	327	11.7	61	2 US-09-027-376-13	Sequence 13, Appl
26	327	11.7	61	2 US-09-094-192-13	Sequence 13, Appl
27	327	11.7	69	2 US-09-268-070-2	Sequence 2, Appl1

28	322	11.5	79	1	US-08-584-671-16	Sequence 16, Appl
29	322	11.5	79	2	US-09-027-376-16	Sequence 16, Appl
30	322	11.5	79	2	US-09-094-192-16	Sequence 16, Appl
31	321	11.5	79	1	US-08-584-671-14	Sequence 14, Appl
32	321	11.5	79	2	US-09-027-376-14	Sequence 14, Appl
33	321	11.5	79	2	US-09-094-192-14	Sequence 14, Appl
34	268.5	9.6	257	2	US-08-596-684F-7	Sequence 7, Appl1
35	239.5	8.6	60	1	US-08-584-671-12	Sequence 12, Appl
36	239.5	8.6	60	2	US-09-027-376-12	Sequence 12, Appl
37	239.5	8.6	60	2	US-09-268-070-4	Sequence 4, Appl1
38	239.5	8.6	60	2	US-09-094-192-12	Sequence 12, Appl
39	238.5	8.6	68	2	US-09-268-070-1	Sequence 1, Appl1
40	211	7.6	40	2	US-09-780-438C-1	Sequence 1, Appl1
41	208	7.5	67	2	US-09-268-070-3	Sequence 3, Appl1
42	201	7.2	38	2	US-09-780-438C-2	Sequence 2, Appl1
43	168.5	6.0	181	2	US-08-948-580-12	Sequence 12, Appl
44	168.5	6.0	181	2	US-08-488-123-12	Sequence 12, Appl
45	163	5.8	514	2	US-10-339-351-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1

Query Match	100.0%	Score 2789;	DB 2;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 9.8e-243;	Mismatches 0;	Indels 0;
Matches 524;	Conservative 0;			Gaps 0;
Qy	1	MYALFLASLILGALAGPVLGLKCEKTRGSVAVCONVKTASDCGAVKHCLQTVNKPVTKS	60	
Db	1	MYALFLASLILGALAGPVLGLKCEKTRGSVAVCONVKTASDCGAVKHCLQTVNKPVTKS	60	
Qy	61	LPDCIKDVVTAAGMDMLKONATEEILVYLEKCDMLPKNNMSASCKEIVDSYLPVLIDI	120	
Db	61	LPDCIKDVVTAAGMDMLKONATEEILVYLEKCDMLPKNNMSASCKEIVDSYLPVLIDI	120	
Qy	121	IKGEMSRPGVCALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPFMANIPLLY	180	
Db	121	IKGEMSRPGVCALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPFMANIPLLY	180	
Qy	181	PQDGPBSKPOKNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI	240	
Db	181	PQDGPBSKPOKNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI	240	
Qy	241	CKNYISQYSBIAIOMMMHMQPKICALVGFCDVEKMPMTLVPAKVASKNVIIPALBLVE	300	
Db	241	CKNYISQYSBIAIOMMMHMQPKICALVGFCDVEKMPMTLVPAKVASKNVIIPALBLVE	300	

QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECEQEV 360
DB 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECEQEV 360
QY 361 VDTYSSILSILLREVSPELVCSMLHLCGSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
DB 361 VDTYSSILSILLREVSPELVCSMLHLCGSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
QY 421 RNLEKNSTKQEILAALEKSGSFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
DB 421 RNLEKNSTKQEILAALEKSGSFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
QY 481 APCSAHKLPGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 524
DB 481 APCSAHKLPGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 524
RESULT 2
US-09-949-016-6272
; Sequence 6272, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6272
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6272
Query Match 100.0%; Score 2789; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 9.8e-243;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
DB 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
QY 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
DB 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
QY 181 PQDGRSKPQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADI 240
DB 181 PQDGRSKPQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKMPQTLVPAKASKNVI PALELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKMPQTLVPAKASKNVI PALELVE 300
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECEQEV 360
DB 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECEQEV 360
QY 361 VDTYSSILSILLREVSPELVCSMLHLCGSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420

DB 361 VDTYSSILSILLREVSPELVCSMLHLCGSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
QY 421 RNLEKNSTKQEILAALEKSGSFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
DB 421 RNLEKNSTKQEILAALEKSGSFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
QY 481 APCSAHKLPGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 524
DB 481 APCSAHKLPGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 524
RESULT 3
US-08-928-074-23
; Sequence 23, Application US/08928074
; Patent No. 6849602
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/928,074
; FILING DATE: 11-SEP-1997
; PRIOR APPLICATION NUMBER: 08/611,307
; FILING DATE: 05-MAR-1996
; PRIOR APPLICATION DATA: PCT/US97/04143
; FILING DATE: 05-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07256/024001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-928-074-23
Query Match 100.0%; Score 2789; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 9.8e-243;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
DB 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
QY 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
DB 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
QY 181 PQDGRSKPQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADI 240

Db 181 PDGPRSKPQKNGDVCQDCIQMTVDIQTAVRTNSTFVQALVEHVKKECDRLGPGVADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKMPMOTLVPKAVSKNVTIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKMPMOTLVPKAVSKNVTIPALELVE 300
Qy 301 PIKKEHVPKSDYVCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSEBCEV 360
Db 301 PIKKEHVPKSDYVCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSEBCEV 360
Qy 361 VDTYGSIIILSEVSPVLVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCCKLVGYLD 420
Db 361 VDTYGSIIILSEVSPVLVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCCKLVGYLD 420
Qy 421 RNLEKNSTQKIEILAELEKGCFLPDYPYQKQCDQFVAEYEPVLTIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTQKIEILAELEKGCFLPDYPYQKQCDQFVAEYEPVLTIEILVEVMDPSFVCLKIG 480
Qy 481 ACPKSAHPLLGTEKCIWGPYSYWCNTTETAACNAVEHCKRHVWN 524
Db 481 ACPKSAHPLLGTEKCIWGPYSYWCNTTETAACNAVEHCKRHVWN 524

RESULT 4

US-09-949-016-8603
; Sequence 8603, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8603
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8603

Query Match 100.0%; Score 2789; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 1e-242;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGALAGPVLGLKECTGSAVWQCNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 12 MYALFLLASLLGALAGPVLGLKECTGSAVWQCNVKTASDCGAVKHCLQTVWVKPTVKS 71
Qy 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTDMLPKPNMGSASCKEIVDSYLPVILDI 120
Db 72 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTDMLPKPNMGSASCKEIVDSYLPVILDI 131
Qy 121 IKGMSRPGVCSALNICSILQKHLAEINHQKLESNKIPELDMTEVVAPFMANIPILLY 180
Db 132 IKGMSRPGVCSALNICSILQKHLAEINHQKLESNKIPELDMTEVVAPFMANIPILLY 191
Qy 181 PDGPRSKPQKNGDVCQDCIQMTVDIQTAVRTNSTFVQALVEHVKKECDRLGPGVADI 240
Db 192 PDGPRSKPQKNGDVCQDCIQMTVDIQTAVRTNSTFVQALVEHVKKECDRLGPGVADI 251
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKMPMOTLVPKAVSKNVTIPALELVE 300
Db 252 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKMPMOTLVPKAVSKNVTIPALELVE 311

Qy 301 PIKKEHVPKSDYVCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSEBCEV 360
Db 312 PIKKEHVPKSDYVCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSEBCEV 371
Qy 361 VDTYGSIIILSEVSPVLVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCCKLVGYLD 420
Db 372 VDTYGSIIILSEVSPVLVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCCKLVGYLD 431
Qy 421 RNLEKNSTQKIEILAELEKGCFLPDYPYQKQCDQFVAEYEPVLTIEILVEVMDPSFVCLKIG 480
Db 432 RNLEKNSTQKIEILAELEKGCFLPDYPYQKQCDQFVAEYEPVLTIEILVEVMDPSFVCLKIG 491
Qy 481 ACPKSAHPLLGTEKCIWGPYSYWCNTTETAACNAVEHCKRHVWN 524
Db 492 ACPKSAHPLLGTEKCIWGPYSYWCNTTETAACNAVEHCKRHVWN 535

RESULT 5

US-08-100-247-2
; Sequence 2, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'Brien.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 99.2%; Score 2766.5; DB 1; Length 523;
Best Local Similarity 99.4%; Pred. No. 1e-240;
Matches 521; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MYALFLLASLLGALAGPVLGLKECTGSAVWQCNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTGSAVWQCNVKTASDCGAVKHCLQTVWVKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTDMLPKPNMGSASCKEIVDSYLPVILDI 120

```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
; US-08-756-031-2

Query Match          99.2%; Score 2766.5; DB 2; Length 523;
Best Local Similarity 99.4%; Pred. No. 1e-240;
Matches 521; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MYALFLASLILGALAGPVLGLKCECTRGSAVWCNQVKTASDCGAVKHCLQTVNKPVTKS 60
Db 1 MYALFLASLILGALAGPVLGLKCECTRGSAVWCNQVKTASDCGAVKHCLQTVNKPVTKS 60
QY 61 LPCDICKDVVTAAGDMLKDNATBEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATBEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVVAPFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVVAPFMANIPLLLY 180
QY 181 PQDGRSKPQPKD-GDVCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
Db 181 PQDGRSKPQPKD-GDVCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALVELVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIIDNNKTEKELTDAPDKMCKSLPKSLSECEQEV 360
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIIDNNKTEKELTDAPDKMCKSLPKSLSECEQEV 360
QY 360 VDTYSSILSLLEEVSPPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Db 360 VDTYSSILSLLEEVSPPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
QY 421 RNLEKNSTKQEIILAEKGCSPDPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEIILAEKGCSPDPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTKEKINGSPSYWCQNTETAACNAVEHCKRHVWN 524
Db 480 ACPSAHKPLLGTKEKINGSPSYWCQNTETAACNAVEHCKRHVWN 523

RESULT 7
US-08-232-513A-3
; Sequence 3, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,031
; FILING DATE: 26-NOV-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'BRIEN.002A
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/100,247
/ FILING DATE: 30-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/DOCKET NUMBER: P-UD 1643
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 535-9001
/ TELEFAX: (619) 535-8949
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 523 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..523
/ OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match      98.9%; Score 2759.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 4.5e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB      1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY      61 LPCDICKDVVTAAGDMLKDNATEEEILVYLKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
DB      61 LPCDICKDVVTAAGDMLKDNATEEEILVYLKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
QY      121 IKGEMSRPGVCSALNLCESLQKHLAEHLNKHQLESNKIPELDMTEVAVPFMANIPILLY 180
DB      121 IKGEMSRPGVCSALNLCESLQKHLAEHLNKHQLESNKIPELDMTEVAVPFMANIPILLY 180
QY      181 PDGPRSKPQPKD-GDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGWADI 240
DB      181 PDGPRSKPQPKD-GDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGWADI 239
QY      241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPAILELVE 300
DB      240 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPAILELVE 299
QY      301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSEECQEV 360
DB      300 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSEECQEV 359
QY      361 VDTYGSLSILLEEVSPELVCSMLHLCSTRLPALTVHVVTQPKDGFCEVCKKLVGYLD 420
DB      360 VDTYGSLSILLEEVSPELVCSMLHLCSTRLPALTVHVVTQPKDGFCEVCKKLVGYLD 419
QY      421 RNLEKNSTQKEILAALEKGCSEFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
DB      420 RNLEKNSTQKEILAALEKGCSEFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 479
QY      481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVECHKRHVN 524
DB      480 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVECHKRHVN 523

RESULT 8
US-08-483-146A-2
/ Sequence 2, Application US/08483146A
/ Patent No. 5696080
/ GENERAL INFORMATION:
/ APPLICANT: O'Brien, John S.
/ APPLICANT: Kishimoto, Yasuo
/ TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
/ TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM
```

```
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Knobbe, Martens, Olson and Bear
/ STREET: 620 Newport Center Blvd. 16th Floor
/ CITY: Newport Beach
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 92660
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/483,146A
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Israel, Ned A.
/ REGISTRATION NUMBER: 29,655
/ REFERENCE/DOCKET NUMBER: MYELOS.002DV1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619-235-8550
/ TELEFAX: 619-235-0176
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 523 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: N-terminal
/ US-08-483-146A-2

Query Match      98.9%; Score 2757.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 6.7e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB      1 MYALFLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY      61 LPCDICKDVVTAAGDMLKDNATEEEILVYLKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
DB      61 LPCDICKDVVTAAGDMLKDNATEEEILVYLKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
QY      121 IKGEMSRPGVCSALNLCESLQKHLAEHLNKHQLESNKIPELDMTEVAVPFMANIPILLY 180
DB      121 IKGEMSRPGVCSALNLCESLQKHLAEHLNKHQLESNKIPELDMTEVAVPFMANIPILLY 180
QY      181 PDGPRSKPQPKD-GDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGWADI 240
DB      181 PDGPRSKPQPKD-GDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLPGWADI 239
QY      241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPAILELVE 300
DB      240 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPAILELVE 299
QY      301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSEECQEV 360
DB      300 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSEECQEV 359
QY      361 VDTYGSLSILLEEVSPELVCSMLHLCSTRLPALTVHVVTQPKDGFCEVCKKLVGYLD 420
DB      360 VDTYGSLSILLEEVSPELVCSMLHLCSTRLPALTVHVVTQPKDGFCEVCKKLVGYLD 419
QY      421 RNLEKNSTQKEILAALEKGCSEFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
DB      420 RNLEKNSTQKEILAALEKGCSEFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 479
```


Query Match 98.9%; Score 2757.5; DB 2; Length 523;
Best Local Similarity 99.2%; Pred. No. 6.7e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MYALFLASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVWVKPTVKS 60

Qy 61 LPCDICKOVVTAAGDMLKONATEBEILVYLKTCMDLKPENMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKOVVTAAGDMLKONATEBEILVYLKTCMDLKPENMSASCKEIVDSYLPVLDI 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180
Db 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLY 180

Qy 181 PQDGRSKPQKONGVQDCIQMWTDIQTAVRTNSTFVQALVHVKEBECDRLPGMADI 240
Db 181 PQDGRSKPQKONGVQDCIQMWTDIQTAVRTNSTFVQALVHVKEBECDRLPGMADI 239

Qy 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDKEMPMQTLVPKAVSKNVPALDELVE 300
Db 240 CKNYISQYSIAIQMMHMQPKKEICALVGFCDKEMPMQTLVPKAVSKNVPALDELVD 299

Qy 301 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMGCKLPKLSSECOEV 360
Db 300 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMGCKLPKLSSECOEV 359

Qy 361 VDTYGSSILSILEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKLVGYLD 420
Db 360 VDTYGSSILSILEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKLVGYLTD 419

Qy 421 RNLEKSTKEIILAALEKGSFLPDPYKQKQDFVARYEPVLEILVEVMDPSFVCLKIG 480
Db 420 RNLEKSTKEIILAALEKGSFLPDPYKQKQDFVARYEPVLEILVEVMDPSFVCLKIG 479

Qy 481 ACPSAHKPLLGTEKINGPSYWCQNTETAQCNVHCKRHHVN 524
Db 480 ACPSAHKPLLGTEKINGPSYWCQNTETAQCNVHCKRHHVN 523

RESULT 11
US-09-352-548-2
; Sequence 2, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; EARLIER FILING DATE: 1999-07-12
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 2
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Saposin B
US-09-352-548-2

Query Match 15.5%; Score 431; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.3e-31;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 GDVQCDCIQMWTDIQTAVRTNSTFVQALVHVKEBECDRLPGMADICKNYISQYSIAIQ 254
Db 1 GDVQCDCIQMWTDIQTAVRTNSTFVQALVHVKEBECDRLPGMADICKNYISQYSIAIQ 60

Qy 255 MMHMQPKKEICALVGFCDK 275
Db 61 MMHMQPKKEICALVGFCDK 81

RESULT 12
US-08-584-671-15
; Sequence 15, Application US/08584671
; Patent No. 5910568
; GENERAL INFORMATION:
; APPLICANT: HAMBERSTEDT, ROY H, BARBATO, GUY F,
; APPLICANT: CRAMER, PALMER
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
; TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
; TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
; ADDRESSEE: STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/584,671
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MONAHAN, THOMAS J
; REGISTRATION NUMBER: 29835
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 814-865-6277
; TELEFAX: 814-865-3591
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: UNKNOWN
US-08-584-671-15

Query Match 15.0%; Score 418.5; DB 1; Length 80;
Best Local Similarity 98.8%; Pred. No. 3.1e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 187
Db 1 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 59

Qy 188 KPQPKONGDVQDCIQMWTDI 208
Db 60 KPQPKONGDVQDCIQMWTDI 80

RESULT 13
US-09-027-376-15
; Sequence 15, Application US/09027376
; Patent No. 6004586
; GENERAL INFORMATION:
; APPLICANT: HAMBERSTEDT, ROY H, BARBATO, GUY F,
; APPLICANT: CRAMER, PALMER
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
; TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
; TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

us-10-801-517-1.ra1

NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-094-192-15

Query Match 14.9%; Score 418.5; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 3.1e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 187
Db 1 PGEVC-ALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 59

Qy 188 KPQKNGDVCDCIQMTDI 208
Db 60 KPQKNGDVCDCIQMTDI 80

RESULT 14
US-09-094-192-15
Sequence 15, Application US/09094192
Patent No. 6103483
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H., BARBATO, GUY F.
APPLICANT: CRAMER, PALMER
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PRO
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESS: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027.376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/584,671
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-027-376-15

Query Match 15.0%; Score 418.5; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 3.1e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 187
Db 1 PGEVC-ALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 59

Qy 188 KPQKNGDVCDCIQMTDI 208
Db 60 KPQKNGDVCDCIQMTDI 80

RESULT 15
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESS: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'Brien.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3

us-10-801-517-1.ra1

NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-094-192-15

Query Match 14.9%; Score 415.5; DB 2; Length 80;
Best Local Similarity 97.5%; Pred. No. 5.7e-30;
Matches 79; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 187
Db 1 PGEVC-ALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGP 59

Qy 188 KPQKNGDVCDCIQMTDI 208
Db 60 KPQKNGDVCDCIQMTDI 80

RESULT 15
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESS: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'Brien.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3

Query Match 14.8%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. NO. 1.2e-29;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCKSLPKSLSEECQEVVDYTGSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCKSLPKSLSEECQEVVDYTGSSILS 60
Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80

Search completed: January 13, 2006, 16:36:34
Job time : 55.6556 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:26:40 ; Search time 99.7682 Seconds
(without alignments)
2194.512 Million cell updates/sec

Title: US-10-801-517-1
Perfect score: 2789
Sequence: 1 MYALFLASLLGAALAGPVL.....NTETAQCNAVEHCKRHVWN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.psp:
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.psp:
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.psp:
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.psp:
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.psp:
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.psp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2789	100.0	524	3	US-09-870-759-60
2	2789	100.0	524	3	US-09-751-708A-60
3	2789	100.0	524	4	US-10-267-502-386
4	2789	100.0	524	4	US-10-408-765A-1207
5	2789	100.0	524	4	US-10-746-442-23
6	2789	100.0	524	4	US-10-428-817A-56
7	2789	100.0	524	5	US-10-801-517-1
8	2789	100.0	524	5	US-10-473-127-1865
9	2789	100.0	524	5	US-10-473-127-1866
10	2789	100.0	524	5	US-10-473-127-1868
11	2789	100.0	524	5	US-10-473-127-1869
12	2789	100.0	524	5	US-10-473-127-1871
13	2789	100.0	524	5	US-10-473-127-1873
14	2789	100.0	524	5	US-10-473-127-1878
15	2789	100.0	524	5	US-10-473-127-1879
16	2789	100.0	524	5	US-10-473-127-1880
17	2789	100.0	524	5	US-10-473-127-1881
18	2789	100.0	524	5	US-10-473-127-1882
19	2789	100.0	524	5	US-10-473-127-2041
20	2789	100.0	524	6	US-11-036-867-23
21	2779	99.6	524	5	US-10-473-127-1874
22	2777.5	99.6	527	3	US-09-870-759-61
23	2777.5	99.6	527	3	US-09-751-708A-61
24	2777.5	99.6	527	4	US-10-060-036-73
25	2777.5	99.6	527	4	US-10-428-817A-57
26	2777.5	99.6	527	5	US-10-473-127-1870
27	2777.5	99.6	527	5	US-10-473-127-1876

28	2772.5	99.4	523	5	US-10-473-127-1877	Sequence 1877, Ap
29	2768	99.2	526	5	US-10-473-127-1875	Sequence 1875, Ap
30	2767.5	99.2	527	5	US-10-473-127-1872	Sequence 1872, Ap
31	2757.5	98.9	523	3	US-09-767-007A-2	Sequence 2, Appl
32	2731	97.9	522	3	US-10-618-281-38	Sequence 38, Appl
33	2449.5	87.8	479	3	US-09-978-418-40	Sequence 40, Appl
34	2449.5	87.8	479	5	US-10-485-231-40	Sequence 40, Appl
35	1969	70.6	554	4	US-10-205-194-176	Sequence 176, App
36	1201	43.1	521	4	US-10-276-162-1	Sequence 1, Appl
37	1201	43.1	521	6	US-11-002-844-1	Sequence 1, Appl
38	1151.5	41.3	227	4	US-10-452-858C-11	Sequence 11, Appl
39	1139.5	40.9	531	3	US-09-833-245-903	Sequence 903, App
40	1125	40.3	210	5	US-10-473-127-1867	Sequence 1867, Ap
41	1093	39.2	209	4	US-10-043-487-340	Sequence 340, App
42	834.5	29.9	362	4	US-10-332-426-8	Sequence 8, Appl
43	552	19.8	953	4	US-10-267-502-385	Sequence 385, App
44	552	19.8	953	6	US-11-097-143-1959	Sequence 1959, Ap
45	545	19.5	241	4	US-10-108-260A-4529	Sequence 4529, Ap

RESULT 1
US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-60

Query Match 100.0%; Score 2789; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MYALFLASLLGAALAGPVLGLKECTRGSAVNCQNVKTASDCGAVKHCLQTWNKPTVKS	60
Db	1	MYALFLASLLGAALAGPVLGLKECTRGSAVNCQNVKTASDCGAVKHCLQTWNKPTVKS	60
QY	61	LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI	120
Db	61	LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI	120
QY	121	IKGMSRPRGEVCSALNLCESLOKHAEHLNKHQLESNKIPELDMTEVAPFMANIPILLY	180
Db	121	IKGMSRPRGEVCSALNLCESLOKHAEHLNKHQLESNKIPELDMTEVAPFMANIPILLY	180
QY	181	PQDGRSKPQKNDGVDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI	240
Db	181	PQDGRSKPQKNDGVDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI	240
QY	241	CKNYISQYSEIAIOMMMHQPKEICALVGFCDEKEMPMQTLVPAKVASKNVIPALELVE	300
Db	241	CKNYISQYSEIAIOMMMHQPKEICALVGFCDEKEMPMQTLVPAKVASKNVIPALELVE	300
QY	301	PIKKEHVPAKSDVYCVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSECOEV	360
Db	301	PIKKEHVPAKSDVYCVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSECOEV	360
QY	361	VDYTGSSILSLLEYSPELVCSMLHLCSTRLPALTVHTVTPQKGGFCVCKLVGYLD	420

ALIGNMENTS

Db 361 VDTYSSSTLSLLEEVSPVCSMLHLCSTRLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Qy 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524

RESULT 2

US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match 100.0%; Score 2789; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
Qy 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQLESNKIPELDMTEVVAPFMANIPLLLY 180
Qy 181 PDGPRSKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Db 181 PDGPRSKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSECCQEV 360
Db 301 PIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSECCQEV 360
Qy 361 VDTYSSSTLSLLEEVSPVCSMLHLCSTRLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Db 361 VDTYSSSTLSLLEEVSPVCSMLHLCSTRLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Qy 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524

RESULT 3

US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match 100.0%; Score 2789; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILI 120
Qy 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQLESNKIPELDMTEVVAPFMANIPLLLY 180
Qy 181 PDGPRSKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Db 181 PDGPRSKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKECDRLGFGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSECCQEV 360
Db 301 PIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKSLSECCQEV 360
Qy 361 VDTYSSSTLSLLEEVSPVCSMLHLCSTRLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Db 361 VDTYSSSTLSLLEEVSPVCSMLHLCSTRLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Qy 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQBIILAALKEGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524

RESULT 4

US-10-408-765A-1207
; Sequence 1207, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

```
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: PASTSEQ for Windows Version 4.0
; SEQ ID NO 1207
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1207

Query Match      100.0%; Score 2789; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGAAALAGPVGLKCECTRGSAVQCQVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLLASLLGAAALAGPVGLKCECTRGSAVQCQVKTASDCGAVKHCLQTVWVKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKONATEEEILVYLEKTCQWLKPKNMSASCKEIVDSYLPVILDI 120
DB 61 LPCDICKDVVTAAGDMLKONATEEEILVYLEKTCQWLKPKNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPALLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPALLY 180

QY 181 PQDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHVHKBECDRGLGPGWADI 240
DB 181 PQDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHVHKBECDRGLGPGWADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300

QY 301 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSESCQEV 360
DB 301 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSESCQEV 360

QY 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420

QY 421 RNLEKSTKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
DB 421 RNLEKSTKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480

QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
DB 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524

RESULT 5
US-10-746-442-23
; Sequence 23, Application US/10746442
; Publication No. US20040121958A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: METHODS FOR ALLEVIATING NEUROPATHIC PAIN
; FILE REFERENCE: 07256/024001
; CURRENT APPLICATION NUMBER: US/10/746,442
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US/08/928,074
; PRIOR FILING DATE: FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 08/611,307
; PRIOR FILING DATE: FILING DATE: 1996-03-05
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: PCT/US97/04143
; PRIOR FILING DATE: FILING DATE: 1996-03-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 23
; LENGTH: 524
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-746-442-23

Query Match      100.0%; Score 2789; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGAAALAGPVGLKCECTRGSAVQCQVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLLASLLGAAALAGPVGLKCECTRGSAVQCQVKTASDCGAVKHCLQTVWVKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKONATEEEILVYLEKTCQWLKPKNMSASCKEIVDSYLPVILDI 120
DB 61 LPCDICKDVVTAAGDMLKONATEEEILVYLEKTCQWLKPKNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPALLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPALLY 180

QY 181 PQDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHVHKBECDRGLGPGWADI 240
DB 181 PQDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHVHKBECDRGLGPGWADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300

QY 301 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSESCQEV 360
DB 301 PIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADPKMCSKLPKSLSESCQEV 360

QY 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420

QY 421 RNLEKSTKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
DB 421 RNLEKSTKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480

QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
DB 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524

RESULT 6
US-10-428-817A-56
; Sequence 56, Application US/10428817A
; Publication No. US20040214783A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 38373-189118
; CURRENT APPLICATION NUMBER: US/10/428,817A
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/378,988
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: US 60/389,366
; PRIOR FILING DATE: 2002-06-15
; PRIOR APPLICATION NUMBER: US 60/406,697
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 60/406,750
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US 60/415,310
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/415,400
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/438,686
; PRIOR FILING DATE: 2003-01-09
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 56
; LENGTH: 524
; TYPE: PRT
```


Db 61 LPDCICKDVVTAAGDMLKDNATEBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQDGRSKPOPKDNGDVCDQCIQWVTDIQTAVRTNSTFVQALVHVHKECDRLGPGWADI 240
Db 181 PQDGRSKPOPKDNGDVCDQCIQWVTDIQTAVRTNSTFVQALVHVHKECDRLGPGWADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVSKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVSKNVIPALELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSEBCEQV 360
Db 301 PIKKEHVPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSEBCEQV 360
Qy 361 VDTYGSLSILLEEVSPPELVCSMLHLCGSTRLPALTUVHTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYGSLSILLEEVSPPELVCSMLHLCGSTRLPALTUVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNSTKQEILAALEKGCFLPDYPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEILAALEKGCFLPDYPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAAQCNVACHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAAQCNVACHCKRHVWN 524

RESULT 9

US-10-473-127-1866
; Sequence 1866, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1866
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1866

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLLASLLGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQDGRSKPOPKDNGDVCDQCIQWVTDIQTAVRTNSTFVQALVHVHKECDRLGPGWADI 240
Db 181 PQDGRSKPOPKDNGDVCDQCIQWVTDIQTAVRTNSTFVQALVHVHKECDRLGPGWADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVSKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVSKNVIPALELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSEBCEQV 360
Db 301 PIKKEHVPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSEBCEQV 360
Qy 361 VDTYGSLSILLEEVSPPELVCSMLHLCGSTRLPALTUVHTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYGSLSILLEEVSPPELVCSMLHLCGSTRLPALTUVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNSTKQEILAALEKGCFLPDYPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEILAALEKGCFLPDYPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAAQCNVACHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAAQCNVACHCKRHVWN 524

RESULT 10

US-10-473-127-1868
; Sequence 1868, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1868
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1868

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLLASLLGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180

121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVAPFMANIPILLY 180
Db
181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240
Qy
181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240
Db
181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240
Qy
241 CKNYISOYSEIATQMMHMQPKIEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALSELVE 300
Db
241 CKNYISOYSEIATQMMHMQPKIEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALSELVE 300
Qy
301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCSKLPKSLSEECQEV 360
Db
301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCSKLPKSLSEECQEV 360
Qy
361 VDTYGSISILLESVSPVLCVSMHLCSGTRLPALTTHVTPQKDGFCFVCKLVGYLD 420
Db
361 VDTYGSISILLESVSPVLCVSMHLCSGTRLPALTTHVTPQKDGFCFVCKLVGYLD 420
Qy
421 RNLEKNSTKOIEILAALEKGCFLPDPYKQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db
421 RNLEKNSTKOIEILAALEKGCFLPDPYKQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Qy
481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524
Db
481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

RESULT 11
US-10-473-127-1869
; Sequence 1869, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1869
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1869

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYAFLLASLLGAALAGPVILGKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVILGKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVAPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVAPFMANIPILLY 180

181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240
Db
181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240
Qy
241 CKNYISOYSEIATQMMHMQPKIEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALSELVE 300
Db
241 CKNYISOYSEIATQMMHMQPKIEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALSELVE 300
Qy
301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCSKLPKSLSEECQEV 360
Db
301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKMCSKLPKSLSEECQEV 360
Qy
361 VDTYGSISILLESVSPVLCVSMHLCSGTRLPALTTHVTPQKDGFCFVCKLVGYLD 420
Db
361 VDTYGSISILLESVSPVLCVSMHLCSGTRLPALTTHVTPQKDGFCFVCKLVGYLD 420
Qy
421 RNLEKNSTKOIEILAALEKGCFLPDPYKQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Db
421 RNLEKNSTKOIEILAALEKGCFLPDPYKQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480
Qy
481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524
Db
481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

RESULT 12
US-10-473-127-1871
; Sequence 1871, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1871
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1871

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYAFLLASLLGAALAGPVILGKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVILGKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVAPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVAPFMANIPILLY 180
Qy 181 PDGPRSKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFVQALVEHVKEECDRLGPGWADI 240

Db 181 PQDGRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLPGGWADI 240
Qy 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVASKNVIPAELVE 300
Db 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVASKNVIPAELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADKMCCKLPKSLSEBCEQV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADKMCCKLPKSLSEBCEQV 360
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKSTQKEILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTQKEILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 13

US-10-473-127-1873
; Sequence 1873, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1873
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-473-127-1873

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLIGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLIGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Qy 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVPFMANIPILLY 180
Qy 181 PQDGRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLPGGWADI 240
Db 181 PQDGRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLPGGWADI 240

Qy 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVASKNVIPAELVE 300
Db 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVASKNVIPAELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADKMCCKLPKSLSEBCEQV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADKMCCKLPKSLSEBCEQV 360
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKSTQKEILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTQKEILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 14

US-10-473-127-1878
; Sequence 1878, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1878
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-473-127-1878

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLIGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLIGALAGPVGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Qy 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKOVVTAAGDMLKDNATEEILVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVPFMANIPILLY 180
Qy 181 PQDGRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLPGGWADI 240
Db 181 PQDGRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLPGGWADI 240
Qy 241 CKNYISQYSIAIQMMHMQPKKEICALVGFCDVEKEMPMTLVPKAVASKNVIPAELVE 300

Db 241 CKNVISQYSEIAIQMMHMQPKETICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKHEVPAKSDVYCEVCEFLVKEVTCLI DNNTKEKEILDADFDMKCSKLPKSLSEECQEV 360
Db 301 PIKKHEVPAKSDVYCEVCEFLVKEVTCLI DNNTKEKEILDADFDMKCSKLPKSLSEECQEV 360
Qy 361 VDTYGSSILSILLBEVSPELVCSMLHLCSGTRLPALT VHVTPQKDGGFCEVCKKLVGYLD 420
Db 361 VDTYGSSILSILLBEVSPELVCSMLHLCSGTRLPALT VHVTPQKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNSTKQRIILAELEKGCFLPDPPYQKQCDQFVAEYEPVLI IELVEMDPSFVCLKIG 480
Db 421 RNLEKNSTKQRIILAELEKGCFLPDPPYQKQCDQFVAEYEPVLI IELVEMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 524

RESULT 15
US-10-473-127-1879
; Sequence 1879, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1879
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1879

Query Match 100.0%; Score 2789; DB 5; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-224;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTWNKPTVKS 60
Qy 61 LPCDICKVVTAGDMKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKVVTAGDMKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGECVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFFMANIPLLIY 180
Db 121 IKGEMSRPGECVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFFMANIPLLIY 180
Qy 181 PQDGRSKPQPKDNGDVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
Db 181 PQDGRSKPQPKDNGDVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADI 240
Qy 241 CKNVISQYSEIAIQMMHMQPKETICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300
Db 241 CKNVISQYSEIAIQMMHMQPKETICALVGFCDVEKEMPQTLVPAKASKNVI PALELVE 300

Qy 301 PIKKHEVPAKSDVYCEVCEFLVKEVTCLI DNNTKEKEILDADFDMKCSKLPKSLSEECQEV 360
Db 301 PIKKHEVPAKSDVYCEVCEFLVKEVTCLI DNNTKEKEILDADFDMKCSKLPKSLSEECQEV 360
Qy 361 VDTYGSSILSILLBEVSPELVCSMLHLCSGTRLPALT VHVTPQKDGGFCEVCKKLVGYLD 420
Db 361 VDTYGSSILSILLBEVSPELVCSMLHLCSGTRLPALT VHVTPQKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNSTKQRIILAELEKGCFLPDPPYQKQCDQFVAEYEPVLI IELVEMDPSFVCLKIG 480
Db 421 RNLEKNSTKQRIILAELEKGCFLPDPPYQKQCDQFVAEYEPVLI IELVEMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 524

Search completed: January 13, 2006, 16:38:41
Job time : 101.768 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006, Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 16:30:30 ; Search time 10.4106 SecGnds
(without alignments)
475.862 Million cell updates/sec

Title: US-10-801-517-1

Perfect score: 2789

Sequence: 1 MYALFLLASILGALAGPVL.....NTETRAQCNVAVHCKRHVWN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 67062 seqs, 9454214 residues

Total number of hits satisfying chosen parameters: 67062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA New:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	6	US-10-821-234-1631
2	113.5	4.1	3803	6	US-10-995-561-773
3	113.5	4.1	3960	6	US-10-995-561-773
4	113.5	4.1	5335	6	US-10-995-561-777
5	113.5	4.1	5406	6	US-10-995-561-774
6	113.5	4.1	5415	6	US-10-995-561-779
7	113.5	4.1	5464	6	US-10-995-561-775
8	113	4.1	1786	7	US-11-196-400-3
9	108.5	3.9	3674	7	US-11-000-463-454
10	107.5	3.9	2036	7	US-11-124-368A-276
11	107.5	3.9	2036	7	US-11-124-368A-280
12	107.5	3.9	2036	7	US-11-124-368A-281
13	107.5	3.9	2044	7	US-11-124-368A-278
14	107.5	3.9	2144	7	US-11-124-368A-277
15	102	3.7	630	7	US-11-196-400-5
16	101	3.6	989	6	US-10-821-234-975
17	98.5	3.5	1404	6	US-10-878-556A-169
18	96.5	3.5	413	6	US-10-821-234-989
19	96	3.4	574	6	US-10-518-341-1
20	96	3.4	574	7	US-11-022-562-214
21	96	3.4	1062	7	US-11-137-465-43
22	95.5	3.4	469	7	US-11-124-368A-321
23	94	3.4	476	6	US-10-878-556A-158
24	94	3.4	585	6	US-10-821-234-875
25	93.5	3.4	782	7	US-11-145-631-9

26	93.5	3.4	930	6	US-10-821-234-1188	Sequence 1188, Ap
27	93.5	3.4	5935	6	US-10-995-561-776	Sequence 776, App
28	93	3.3	3002	6	US-10-821-234-916	Sequence 916, App
29	92.5	3.3	754	6	US-10-467-962B-63	Sequence 63, Appli
30	92.5	3.3	1163	7	US-11-044-899-2	Sequence 2, Appli
31	92.5	3.3	1163	7	US-11-044-899-30	Sequence 30, Appli
32	90.5	3.2	459	6	US-10-763-712A-70	Sequence 70, Appli
33	90.5	3.2	503	6	US-10-878-556A-2	Sequence 2, Appli
34	90.5	3.2	537	6	US-10-821-234-1427	Sequence 1427, Ap
35	90.5	3.2	741	7	US-11-191-374-18	Sequence 18, Appli
36	90.5	3.2	741	7	US-11-191-375-18	Sequence 18, Appli
37	90.5	3.2	741	7	US-11-191-588-18	Sequence 18, Appli
38	90.5	3.2	1586	6	US-10-821-234-901	Sequence 901, App
39	90.5	3.2	2048	7	US-11-116-939-6	Sequence 6, Appli
40	89.5	3.2	420	6	US-10-131-826A-230	Sequence 230, App
41	89.5	3.2	618	6	US-10-523-912-4	Sequence 4, Appli
42	89.5	3.2	1933	6	US-10-523-912-2	Sequence 2, Appli
43	89.5	3.2	5024	6	US-10-793-626-2964	Sequence 2964, Ap
44	89	3.2	556	6	US-10-995-561-766	Sequence 766, App
45	89	3.2	556	6	US-10-995-561-767	Sequence 767, App

ALIGNMENTS

RESULT 1
US-10-821-234-1631
; Sequence 1631, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang Y. Tom
; TITLE OF INVENTION: Methods for diagnosis and treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1631
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1631

Query Match	100.0%	Score	2789;	DB	6;	Length	524;
Best Local Similarity	100.0%	Pred. No.	4e-211;				
Matches	524;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MYALFLLASILGALAGPVLGLKCTRGSAVWQNVKTASDCCGAVKHCLQTVWVKPTVKS	60				
Db	1	MYALFLLASILGALAGPVLGLKCTRGSAVWQNVKTASDCCGAVKHCLQTVWVKPTVKS	60				
Qy	61	LPDCICDVVTRAGDMLKDNATEEILVYLEKTCDMLPKNMSASCKEIVDSYLPVLIDI	120				
Db	61	LPDCICDVVTRAGDMLKDNATEEILVYLEKTCDMLPKNMSASCKEIVDSYLPVLIDI	120				
Qy	121	IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLY	180				
Db	121	IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLY	180				
Qy	181	PDGPRSKPQKNGDVQCDICIQMTVDIQTAVRTNSTFVQALVHVHKEECDRILGPGNADI	240				
Db	181	PDGPRSKPQKNGDVQCDICIQMTVDIQTAVRTNSTFVQALVHVHKEECDRILGPGNADI	240				
Qy	241	CKNYISQYSEIATOMMMHMQPKEICALVGFCDEKMPQTLVPAKVASKNVIPALELVE	300				
Db	241	CKNYISQYSEIATOMMMHMQPKEICALVGFCDEKMPQTLVPAKVASKNVIPALELVE	300				
Qy	301	PIKHEVPAKSDVYCEVCEFLVKEVTYGLIDNNKTEKELIDAFDKWCSKLPKSLSECEQV	360				

```

Db      301  |||||||PIKKEVPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDADKWCCKLPKSLSECOEV 360
Qy      361  VDTYSSILSLLEEVSELCVSMHLCSGRLPALTVHVTOPKGGFCVCKKLVGYLD 420
Db      361  VDTYSSILSLLEEVSELCVSMHLCSGRLPALTVHVTOPKGGFCVCKKLVGYLD 420
Qy      421  RNLEKNSKQIBLAALGKCSFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Db      421  RNLEKNSKQIBLAALGKCSFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Qy      481  ACPSAHKELLGTEKICWGPSYWCNTETAACQNAVEHCRRHVN 524
Db      481  ACPSAHKELLGTEKICWGPSYWCNTETAACQNAVEHCRRHVN 524

```

```

RESULT* 2
US-10-995-561-773
; Sequence 773, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773
; LENGTH: 3803
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-773

```

```

Query Match      4.1%; Score 113.5; DB 6; Length 3803;
Best Local Similarity 16.3%; Pred. No. 1.4;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;

Qy      71  TAAGDMLKDNATEEE-----ILVYLEKTCDWLPKPNMSASCKEIVDSYLPVIL 118
Db      1295 TAQEMMKPGQAEDSRVLSQLSQOTAFABIERNTQTKLDQ-----CQKFSQQYSTIVK 1348
Qy      119  D-----IIKEMSRPGEVCSALNLCESLQKHLAEL-----NHQKQL-- 154
Db      1349 DYELQMTYKAFVESQOKSPGRRRLSSDAITQEFMDLTRYALTALVTLTQHVKYISD 1408
Qy      155  -----ESNKIPE---LDMTEVVAPFMANIPLLLYPQDGPSPKQPONGDV----- 197
Db      1409 ALRLEEEKVVEEKEQHEVKEKLGWVSTLARNTQKATSETKSTDIKAILQQ 1468
Qy      198  --CQDCIQMVTDIQTAVRTNSTFVQALVEH-----VKECDRLPGMADICKNY 244
Db      1469 VLSELTTKKEQVSEAIKTSQIF---LAKHGKLSKEKKQISEQLNALNKAYHDLCDGS 1525
Qy      245  ISQYSEIAIQM-----MMHMQPKIECALVGFCDVEKEMPQTLVPAKVASK 290
Db      1526 ANOLOQSQLAHQTEQKTLQKQNTCHQLEDLCSWYGOAERA-----LAGH 1573
Qy      291  NVIPALELVEPIKKEVPAK---SDVYCEVCEP---LVKEVTKLIDNNKTE---KEILDAF 342
Db      1574 QGRITQODLSALQKNQSDLKLDQDDIQNRATSPATVVKDIEGFMEENQTKLSPRELTALR 1633
Qy      343  DKM--CSKLPKSLSECOEVVDYTGSSILSILEEVSPELVCSMLHLCSGTRLPALTAVHV 400
Db      1634 EKLHQAKEQYALQEBETVAQKELEAVTSALQOETEKSKAAKEL--AENKKIDALLDWV 1692
Qy      401  TQPKDGGFCVCKKLVGYLDRLNLEKNSKQIBLAALGKGC-----SFLPDPYQKQ 450
Db      1693 TSVGSSG-----GQLLTNLP--GMEQLSGASLEKGLDITTDGYMGVNOAPEKLDKQ 1741
Qy      451  CQDQVAYEYEPVL 462

```

```

Db      1742 CEMMKARHQELL 1753

RESULT 3
US-10-995-561-771
; Sequence 771, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771
; LENGTH: 3960
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-771

```

```

Query Match      4.1%; Score 113.5; DB 6; Length 3960;
Best Local Similarity 16.3%; Pred. No. 1.4;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;

Qy      71  TAAGDMLKDNATEEE-----ILVYLEKTCDWLPKPNMSASCKEIVDSYLPVIL 118
Db      1295 TAQEMMKPGQAEDSRVLSQLSQOTAFABIERNTQTKLDQ-----CQKFSQQYSTIVK 1348
Qy      119  D-----IIKEMSRPGEVCSALNLCESLQKHLAEL-----NHQKQL-- 154
Db      1349 DYELQMTYKAFVESQOKSPGRRRLSSDAITQEFMDLTRYALTALVTLTQHVKYISD 1408
Qy      155  -----PSNKIPE---LDMTEVVAPFMANIPLLLYPQDGPSPKQPONGDV----- 197
Db      1409 ALRLEEEKVVEEKEQHEVKEKLGWVSTLARNTQKATSETKSTDIKAILQQ 1468
Qy      198  --CQDCIQMVTDIQTAVRTNSTFVQALVEH-----VKECDRLPGMADICKNY 244
Db      1469 VLSELTTKKEQVSEAIKTSQIF---LAKHGKLSKEKKQISEQLNALNKAYHDLCDGS 1525
Qy      245  ISQYSEIAIQM-----MMHMQPKIECALVGFCDVEKEMPQTLVPAKVASK 290
Db      1526 ANOLOQSQLAHQTEQKTLQKQNTCHQLEDLCSWYGOAERA-----LAGH 1573
Qy      291  NVIPALELVEPIKKEVPAK---SDVYCEVCEP---LVKEVTKLIDNNKTE---KEILDAF 342
Db      1574 QGRITQODLSALQKNQSDLKLDQDDIQNRATSPATVVKDIEGFMEENQTKLSPRELTALR 1633
Qy      343  DKM--CSKLPKSLSECOEVVDYTGSSILSILEEVSPELVCSMLHLCSGTRLPALTAVHV 400
Db      1634 EKLHQAKEQYALQEBETVAQKELEAVTSALQOETEKSKAAKEL--AENKKIDALLDWV 1692
Qy      401  TQPKDGGFCVCKKLVGYLDRLNLEKNSKQIBLAALGKGC-----SFLPDPYQKQ 450
Db      1693 TSVGSSG-----GQLLTNLP--GMEQLSGASLEKGLDITTDGYMGVNOAPEKLDKQ 1741
Qy      451  CQDQVAYEYEPVL 462

```

```

RESULT 4
US-10-995-561-777
; Sequence 777, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

```

;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001559
;; CURRENT APPLICATION NUMBER: US/10/995,561
;; CURRENT FILING DATE: 2004-11-24
;; NUMBER OF SEQ ID NOS: 85702
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 777
;; LENGTH: 5335
;; TYPE: PRN
;; ORGANISM: Homo sapiens
US-10-995-561-777

Query Match 4.1%; Score 113.5; DB 6; Length 5335;
Best Local Similarity 16.3%; Pred. No. 2.1;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;
Qy 71 TAAGMLKDNATEEE-----ILVLEKTCWLPKPNMSASCKEIVDSYLPVIL 118
Db 1203 TAQEMMKPGQAEBSRVLSQSQQTALFAIERNTKLDQ-----CQFSQOYSTIVK 1256
Qy 119 D-----IIKGEMSRPGVCSALNLCESLQKHLAE-----NHQKQL-- 154
Db 1257 DYELQMTYKAFVESQKSPGKRRLSSSDAITQEFMDLTRYALTALVLTQHVKYISD 1316
Qy 155 -----ESNKIPE---LDMTEVAPFMANIPLLYPQDGPSPKQPKONGDV----- 197
Db 1317 ALRLEEBEKVVEBKQHVKEKLLGWSTLARNTQKATSSSTKSTEDIEKAILREQ 1376
Qy 198 --CODCIQWTDIQTAVTNTSTFVQALVEH-----VKECDRLGPGMADICKNY 244
Db 1377 VLSELTTKKQVSEAIKTSQIF---LAKHGHLSEKKEKQISEQNALNKAHYHDLCDGS 1433
Qy 245 ISQYSEIAIOM-----MMHQPKETCALVGFCDVCKEMPQTLVPAKVASK 290
Db 1434 ANQLQQLQSQAHTQETKLOKQNTCHQQLDLCWVQAE-----LAGH 1481
Qy 291 NVIPALBELVPIKHEVPAK---SDVYCEVCEP---LVKEVTKLIDNNKTE---KEILDAF 342
Db 1482 QGRITQDLSALQKNSDLQDDIQNRATSFATVVKDIEGFMEENQTKLSPRELTALR 1541
Qy 343 DKM--CSKLPKSLSECEVVDVTGSSILSLLEEVSPVCLVCSMLHLCGTRLPALTIVH 400
Db 1542 EKLHQAQEYEQALQETRVAKQELAEAVTSALQOETESKKAAKEL-AENKKIKDALLDWV 1600
Qy 401 TQPKDGGCFCEVCKKLVGLDRLNLEKNTKQELIAALEKGC-----SFLPDPYQKQ 450
Db 1601 TSVGSSG-----GQLLTNLP--GMEQLSGASLEKGLDITTDGYNVGNQAPKLDKQ 1649
Qy 451 CDQFVAEYEPVL 462
Db 1650 CENMKARHQBELL 1661

RESULT 5
US-10-995-561-774
;; Sequence 774, Application US/10995561
;; Publication No. US20050272054A1
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele et al.
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001559
;; CURRENT APPLICATION NUMBER: US/10/995,561
;; CURRENT FILING DATE: 2004-11-24
;; NUMBER OF SEQ ID NOS: 85702
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 774
;; LENGTH: 5406
;; TYPE: PRN
;; ORGANISM: Homo sapiens
US-10-995-561-774

Query Match 4.1%; Score 113.5; DB 6; Length 5406;
Best Local Similarity 16.3%; Pred. No. 2.1;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;
Qy 71 TAAGMLKDNATEEE-----ILVLEKTCWLPKPNMSASCKEIVDSYLPVIL 118
Db 1295 TAQEMMKPGQAEBSRVLSQSQQTALFAIERNTKLDQ-----CQFSQOYSTIVK 1348
Qy 119 D-----IIKGEMSRPGVCSALNLCESLQKHLAE-----NHQKQL-- 154
Db 1349 DYELQMTYKAFVESQKSPGKRRLSSSDAITQEFMDLTRYALTALVLTQHVKYISD 1408
Qy 155 -----ESNKIPE---LDMTEVAPFMANIPLLYPQDGPSPKQPKONGDV----- 197
Db 1409 ALRLEEBEKVVEBKQHVKEKLLGWSTLARNTQKATSSSTKSTEDIEKAILREQ 1468
Qy 198 --CODCIQWTDIQTAVTNTSTFVQALVEH-----VKECDRLGPGMADICKNY 244
Db 1469 VLSELTTKKQVSEAIKTSQIF---LAKHGHLSEKKEKQISEQNALNKAHYHDLCDGS 1525
Qy 245 ISQYSEIAIOM-----MMHQPKETCALVGFCDVCKEMPQTLVPAKVASK 290
Db 1526 ANQLQQLQSQAHTQETKLOKQNTCHQQLDLCWVQAE-----LAGH 1573
Qy 291 NVIPALBELVPIKHEVPAK---SDVYCEVCEP---LVKEVTKLIDNNKTE---KEILDAF 342
Db 1574 QGRITQDLSALQKNSDLQDDIQNRATSFATVVKDIEGFMEENQTKLSPRELTALR 1633
Qy 343 DKM--CSKLPKSLSECEVVDVTGSSILSLLEEVSPVCLVCSMLHLCGTRLPALTIVH 400
Db 1634 EKLHQAQEYEQALQETRVAKQELAEAVTSALQOETESKKAAKEL-AENKKIKDALLDWV 1692
Qy 401 TQPKDGGCFCEVCKKLVGLDRLNLEKNTKQELIAALEKGC-----SFLPDPYQKQ 450
Db 1693 TSVGSSG-----GQLLTNLP--GMEQLSGASLEKGLDITTDGYNVGNQAPKLDKQ 1741
Qy 451 CDQFVAEYEPVL 462
Db 1742 CENMKARHQBELL 1753

RESULT 6
US-10-995-561-779
;; Sequence 779, Application US/10995561
;; Publication No. US20050272054A1
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele et al.
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001559
;; CURRENT APPLICATION NUMBER: US/10/995,561
;; CURRENT FILING DATE: 2004-11-24
;; NUMBER OF SEQ ID NOS: 85702
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 779
;; LENGTH: 5415
;; TYPE: PRN
;; ORGANISM: Homo sapiens
US-10-995-561-779

Query Match 4.1%; Score 113.5; DB 6; Length 5415;
Best Local Similarity 16.3%; Pred. No. 2.1;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;
Qy 71 TAAGMLKDNATEEE-----ILVLEKTCWLPKPNMSASCKEIVDSYLPVIL 118
Db 1295 TAQEMMKPGQAEBSRVLSQSQQTALFAIERNTKLDQ-----CQFSQOYSTIVK 1348
Qy 119 D-----IIKGEMSRPGVCSALNLCESLQKHLAE-----NHQKQL-- 154
Db 1349 DYELQMTYKAFVESQKSPGKRRLSSSDAITQEFMDLTRYALTALVLTQHVKYISD 1408


```

QY      231 NVIPALELVEPIKKHEVPAK---SDVIYCEVCEF--LVAEVINADLNKNKL
DB      1574 QGRFTQQDLSALQKNQSDLQODDIQNRATSPFATVWKDIEGFEENOTPKLSPRELTALR 1633
QY      343 DKM---CSKLPSLSBECQEVVDVTYGSSILSILEEVSPELVCSMLHLCSGTRLPALTVHV 400
DB      1634 EKLHQAKEQYEALQEETRVAKELBEAVTSALQOETEKSAAKEL-AENKKIDALLDMW 1692
QY      401 TQPDGGFCFCEVKCLVGYLDRNLKKNSTKQETLALEKGC-----SFLPDPYPQKQ 450
DB      1693 TSVSGSG-----GQLLTNLP--GMSQLSGASLEKGALDTTDDGVGMVGNQAPEKLDKQ 1741
QY      451 CDQFVAEYEPVL 462
DB      1742 CEMMKARHOELL 1753

RESULT 8
US-11-196-400-3
; Sequence 3, Application US/11196400
; Publication No. US20050287166A1
; GENERAL INFORMATION:
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: WALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USDIV
; CURRENT APPLICATION NUMBER: US/11/196,400
; PRIOR FILING DATE: 2005-08-04
; PRIOR APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,462
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-11-196-400-3

Query Match          4.1%; Score 113; DB 7; Length 1786;
Best Local Similarity 19.5%; Pred. No. 0.56;
Matches 94; Conservative 97; Mismatches 169; Indels 122; Gaps 27;

QY      56 PTWKSPLC-DICKDVTTAGDMLKDNAMEEILLYLEKTCDMLPKPMNSASCBE-IVDSY 113
DB      737 PTVEEIVAPSVAPSVEESVEEN-VESVAENVEESV----AENVEESVAENVEESV 791
QY      114 LPVLTDIIGKEMSRPGVCFSALNLCBSLQKHAB--LNHQKQLESNKIPELDMTEVVAP 170
DB      792 APTVEEIVAPSVEE-----SVAPSVEESAENVATNLSDNLNLGLGIEETEEKDS 843
QY      171 FWANIPLLXYPDGFRSKPQDKONGDVQCDCIQMTDI-----QTAVRTNSTFVALVEH 225
DB      844 ILNEI-----EEYKEN-----VVTTLLENVEETTAESTVTF-SNILEE 880
QY      226 VKREC-----DRLGPGMADICKNVISQYSIAIQMMHMOPKEICALVGPCCDEKEMPQTL 282
DB      881 IQENTITNDTITEEKLEEHENVLSSALE---NTQSEEEKKEVIDVI---EEVKEEVATTLL 934
QY      283 V-----PAKVASKNPVPALELVEPIKGEHPAKSDYVCEFL-----VKEBTK 327
DB      935 IETVEQAEKKSANTI--TEIFPENLENAVESNENV-AENLEKLNFTVFTVLDKVEETVE 991
QY      328 L-----IDNNKTBKIL-DAFDK-----MCSKLPKSLSEBCQEVVDVTYGSSILSI 371
DB      992 ISGESLENMEMDKAFPEFIDFNKGTQENLLTGMPRSIETSIVIYQSEEKVDLENVVSII 1051
QY      372 L---LEEVSPELVCSMLHLCSGTRL-PALTAVHTVTPQKDGFCFCEVKCLVGYLDRNLKKNST 428

```

```

155 -----ESNKIPB---LDMTEVVAPFMANIPLLYPQDGRSKPQPKDNGDV----- 197
1409 ALRRLBEEBKVEEKEHVEKVELLGMVSTLARNTQKATSETKSTEDIEKALEEQ 1468
198 --QDCIQWTDITQAVRTNSTPQALVEH-----VKECDRLPGMGADICKNY 244
1469 VLSEELTTKKEQVSEAIKTSQIF---LAKGHKLSKEKKQISEQLNALNKAYHDLCDGS 1525
245 ISQVSEIAIOM-----NMHMQPKBICALVGFCDVEKEMPMOTLVPKAVASK 290
1526 ANOLOQLQSLAHQTEQTLQKQNTCHQLEDLCSWVGOAERA-----LAGH 1573
291 NVIPALSELVEPIKKEHVPK---SDVYCEVCEF--LVKEVTKLIDNNKTE--KEILDAF 342
1574 QGRTOQDLSALQNSQDLQLODDIQNRATSFATVVKDIEGFERNQTKLSPRELTALR 1633
343 DKM--CSKLPKSLSECEQVVDTYGSSILSLLEEVSPELVCSMLHLCSGTRLPALTVHV 400
1634 EKLHQAKQYEQALQEBETRAQKELEEAVTSALQOETEKSAAKEL--AENKKKIDALLDWV 1692
401 TOPKXGGCFCEVCKLVGLYLDRLNLEKNSTKQETLAALEKGC-----SFLPDPYQOK 450
1693 TSVGSSG-----GOLLTNLP--GMEQLSGASLEKGAULTDTPGYMGVNGAPEKLDKQ 1741
451 CDOFVAEYEPVL 462
1742 CEMMKARHQELL 1753

RESULT 7
US-10-995-561-775
; Sequence 775, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 5464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-775

Query Match 4.1%; Score 113.5; DB 6; Length 5464;
Best Local Similarity 16.3%; Pred. No. 2.2;
Matches 80; Conservative 88; Mismatches 191; Indels 133; Gaps 19;

QY 71 TAAGDMLKDNATPEE-----ILVLEKTCMDLPRKNMSASCKEIVDSYLPVIL 118
Db 1295 TAAQEMKPGQAEKDSRVLSQSLQOOTALFBIERNQTKLDQ-----CQKSEQYSTIVK 1348
QY 119 D-----IKGEMSPGVCNALCSLQKHLAE-----NHQKQL-- 154
Db 1349 DYELQLMTYKAFVSEQOKSPGRRRMLSSDAITQFMDLTRYTALTALVTLTTHQVKYISD 1408
QY 155 -----ESNKIPB---LDMTEVVAPFMANIPLLYPQDGRSKPQPKDNGDV----- 197
Db 1409 ALRRLBEEBKVEEKEHVEKVELLGMVSTLARNTQKATSETKSTEDIEKALEEQ 1468
QY 198 --QDCIQWTDITQAVRTNSTPQALVEH-----VKECDRLPGMGADICKNY 244
Db 1469 VLSEELTTKKEQVSEAIKTSQIF---LAKGHKLSKEKKQISEQLNALNKAYHDLCDGS 1525
245 ISQVSEIAIOM-----NMHMQPKBICALVGFCDVEKEMPMOTLVPKAVASK 290
1526 ANOLOQLQSLAHQTEQTLQKQNTCHQLEDLCSWVGOAERA-----LAGH 1573

```


Wed Jan 18 11:58:40 2006

```

995 HLT-----LGNKNEFSVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 994
283 VPAKVASKNVIPALBELVPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325
995 -----QITALEKITKPPFAAISDEKVOQKLLRMLFDLLVNCKNSHCAQTSSV 1042
326 TKLIDNNKTEKEI-LDAFD-----KMSCKLPKSLSECOEVDVDTYGS--SILSIL 372
1043 PKGISVNAEQVRIELEPPDKAKPLGTVOQKRQRKQKKSQDLESQVSGGSYMQRVTLI 1102
373 LEEV-----SPB-LVCSMLHLCSTRLPALTVHTVTPQKGGFCFCEVCKLVGYLDRLNE 424
1103 LELLQHKKLSRQILVPTLFNLLSRCLEPL-----POEQG-----NME 1141
425 KNSTQKEILAALEKGCSEFLPDYQKQDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPS 484
1142 --YTKQILSLCLNLCQKL-SPDGGKIPKIDLEKFNVELIVQ-----CIRLSEMPQ 1191
485 AHKP---LLGT 492
1192 THHALLLIGT 1202

RESULT 12
US-11-124-368A-281
; Sequence 281, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281
; LENGTH: 2036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-281

Query Match 3.9%; Score 107.5; DB 7; Length 2036;
Best Local Similarity 20.0%; Pred. No. 1.8; Indels 159; Gaps 27;
Matches 110; Conservative 91; Mismatches 191;

QY 1 MYAFLLASLL---GAALAGPV-LGLKECTRGSAVCONVKTASDCGAVKHCLQTVNKP 56
752 VFQLFKCSVLMTYSSLSNPLNCVKTIVLTQALYVGCAMLSSQKTCQKHQLASI-SSP 810
57 TVKSLPCDI---CKDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSY 113
811 VVTSLLINLGSFVKEVRRAA-----IQCLQALSGV 840
114 LPVILDIKGBMSRPGVCS-ALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFM 172
841 ASPFYLIIDHLISKAEIITSDAAVYIQLATLFEELQREKGLKSH-----QKLSL 892
173 ANIPLLYPQDGRPSKPOPK-----DNGD-VCQDCIQMVTDIQTAVRTNSTFV---QALVE 224
893 KNLSCVYSCPSYIAKMLKVLQGVNGEVLSQLPMAEQLEKIQKEPTAVLKDEAMVL 952
225 HVKECDRLGPGMADICKNYISQYSEIAIQMMHMQPKEICALVGFCDCEVKEM--PMQTL 282
953 HLT-----LGKYNFVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 994
283 VPAKVASKNVIPALBELVPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325
995 -----QITALEKITKPPFAAISDEKVOQKLLRMLFDLLVNCKNSHCAQTSSV 1042

```

```

953 HLT-----LGNKNEFSVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 994
283 VPAKVASKNVIPALBELVPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325
995 -----QITALEKITKPPFAAISDEKVOQKLLRMLFDLLVNCKNSHCAQTSSV 1042
326 TKLIDNNKTEKEI-LDAFD-----KMSCKLPKSLSECOEVDVDTYGS--SILSIL 372
1043 PKGISVNAEQVRIELEPPDKAKPLGTVOQKRQRKQKKSQDLESQVSGGSYMQRVTLI 1102
373 LEEV-----SPB-LVCSMLHLCSTRLPALTVHTVTPQKGGFCFCEVCKLVGYLDRLNE 424
1103 LELLQHKKLSRQILVPTLFNLLSRCLEPL-----POEQG-----NME 1141
425 KNSTQKEILAALEKGCSEFLPDYQKQDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPS 484
1142 --YTKQILSLCLNLCQKL-SPDGGKIPKIDLEKFNVELIVQ-----CIRLSEMPQ 1191
485 AHKP---LLGT 492
1192 THHALLLIGT 1202

RESULT 11
US-11-124-368A-280
; Sequence 280, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 280
; LENGTH: 2036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-280

Query Match 3.9%; Score 107.5; DB 7; Length 2036;
Best Local Similarity 20.0%; Pred. No. 1.8; Indels 159; Gaps 27;
Matches 110; Conservative 91; Mismatches 191;

QY 1 MYAFLLASLL---GAALAGPV-LGLKECTRGSAVCONVKTASDCGAVKHCLQTVNKP 56
752 VFQLFKCSVLMTYSSLSNPLNCVKTIVLTQALYVGCAMLSSQKTCQKHQLASI-SSP 810
57 TVKSLPCDI---CKDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSY 113
811 VVTSLLINLGSFVKEVRRAA-----IQCLQALSGV 840
114 LPVILDIKGBMSRPGVCS-ALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFM 172
841 ASPFYLIIDHLISKAEIITSDAAVYIQLATLFEELQREKGLKSH-----QKLSL 892
173 ANIPLLYPQDGRPSKPOPK-----DNGD-VCQDCIQMVTDIQTAVRTNSTFV---QALVE 224
893 KNLSCVYSCPSYIAKMLKVLQGVNGEVLSQLPMAEQLEKIQKEPTAVLKDEAMVL 952
225 HVKECDRLGPGMADICKNYISQYSEIAIQMMHMQPKEICALVGFCDCEVKEM--PMQTL 282
953 HLT-----LGKYNFVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 994
283 VPAKVASKNVIPALBELVPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325

```

QY 326 TKLIDNNKTEKEI-LDAFD-----KMCCKLPKSLSEBCEQVVDVYGS--SILSIL 372
Db 1043 FKGISVNAEQVRIELEPPDKAKPLGTQVQRRQKMQQKKSQDLESQVGGSYWQVRVTLI 1102
QY 373 LEEV-----SPB-LVCSMLHLCSTRLPALTVHVTPQKDGFCFCEVCKLVGYLDNRLE 424
Db 1103 LELLQHKKKLRSFQILVPTLFNLLSRCLEPL-----PQEQ-----NME 1141
QY 425 KNSTKQBIILAALKEGCSFLPDPYQKQDOFVAEYEPVLBIELVEMDPSFVCLKIGACPS 484
Db 1142 --YTKQILISCLNLCQKL-SPDGGKIPKDIIDBEKFNVELIVQ-----CIRLSEMPQ 1191
QY 485 AHKP---LLGT 492
Db 1192 THHALLLLGT 1202

RESULT 13

US-11-124-368A-278
; Sequence 278, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 2044
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-278

Query Match 3.9%; Score 107.5; DB 7; Length 2044;

Best Local Similarity 20.0%; Pred. No. 1.8;
Matches 110; Conservative 91; Mismatches 191; Indels 159; Gaps 27;

QY 1 MYALFLASLL---GAALAGPV-LGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWVNP 56
Db 760 VFQLFKFCVLTWYTGSSLNPLNCSVKTVLQALYVGCAMLSSQKTOCKHQLASI-SSP 818
QY 57 TVKSLPCDI---CKDVVTAAGDMLKDNATEEBEILVYLEKTDWLKPKNMSASCKEIVDSY 113
Db 819 VVTSLLINLGSVPKEVRRAA-----IQCLQALSGV 848
QY 114 LPVILDIKEMSRPGEVCS-ALNLCESLQKHLAEHLNHOLESNKIPBLDMTEVVAPFM 172
Db 849 ASPFYLIIDHLISKABEITSDAAYVIQDLATLFEELQREKLLKSH-----QKLSSETL 900
QY 173 ANIPLLLYPDGPRSKPQPK-----DNGD-VCODCTQMVTDIQTAVRTNSTFV---QALVE 224
Db 901 KNLLSCVYSCPSYIAKDLMKVLQGVNGENVLSQLLPMAEQLEKIQKEPTAVLKDQEAAML 960
QY 225 HVKEBCDRGLGPGMADICKNVIQSYSEIAIQMMHMQPKIEICALVGFCDEVKEM--PMOTL 282
Db 961 HLT-----LGKYNFVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 1002
QY 283 VPAKASKNVIPALELVEPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325
Db 1003 -----QITALEKITPPFAAISDEKVVQKLLRMLFDLLVNCNKSCHCAQTSSV 1050
QY 326 TKLIDNNKTEKEI-LDAFD-----KMCCKLPKSLSEBCEQVVDVYGS--SILSIL 372
Db 1051 FKGISVNAEQVRIELEPPDKAKPLGTQVQRRQKMQQKKSQDLESQVGGSYWQVRVTLI 1110

QY 373 LEEV-----SPB-LVCSMLHLCSTRLPALTVHVTPQKDGFCFCEVCKLVGYLDNRLE 424
Db 1111 LELLQHKKKLRSFQILVPTLFNLLSRCLEPL-----PQEQ-----NME 1149
QY 425 KNSTKQBIILAALKEGCSFLPDPYQKQDOFVAEYEPVLBIELVEMDPSFVCLKIGACPS 484
Db 1150 --YTKQILISCLNLCQKL-SPDGGKIPKDIIDBEKFNVELIVQ-----CIRLSEMPQ 1199
QY 485 AHKP---LLGT 492
Db 1200 THHALLLLGT 1210

RESULT 14

US-11-124-368A-277
; Sequence 277, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-277

Query Match 3.9%; Score 107.5; DB 7; Length 2144;

Best Local Similarity 20.0%; Pred. No. 1.9;
Matches 110; Conservative 91; Mismatches 191; Indels 159; Gaps 27;

QY 1 MYALFLASLL---GAALAGPV-LGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWVNP 56
Db 860 VFQLFKFCVLTWYTGSSLNPLNCSVKTVLQALYVGCAMLSSQKTOCKHQLASI-SSP 918
QY 57 TVKSLPCDI---CKDVVTAAGDMLKDNATEEBEILVYLEKTDWLKPKNMSASCKEIVDSY 113
Db 919 VVTSLLINLGSVPKEVRRAA-----IQCLQALSGV 948
QY 114 LPVILDIKEMSRPGEVCS-ALNLCESLQKHLAEHLNHOLESNKIPBLDMTEVVAPFM 172
Db 949 ASPFYLIIDHLISKABEITSDAAYVIQDLATLFEELQREKLLKSH-----QKLSSETL 1000
QY 173 ANIPLLLYPDGPRSKPQPK-----DNGD-VCODCTQMVTDIQTAVRTNSTFV---QALVE 224
Db 1001 KNLLSCVYSCPSYIAKDLMKVLQGVNGENVLSQLLPMAEQLEKIQKEPTAVLKDQEAAML 1060
QY 225 HVKEBCDRGLGPGMADICKNVIQSYSEIAIQMMHMQPKIEICALVGFCDEVKEM--PMOTL 282
Db 1061 HLT-----LGKYNFVS-LLNEDPKSLDIFIKAVHTTKELYAGMPTI 1102
QY 283 VPAKASKNVIPALELVEPIKKEHVEPAKSD-----VYCE--VCEFLVKEV 325
Db 1103 -----QITALEKITPPFAAISDEKVVQKLLRMLFDLLVNCNKSCHCAQTSSV 1150
QY 326 TKLIDNNKTEKEI-LDAFD-----KMCCKLPKSLSEBCEQVVDVYGS--SILSIL 372
Db 1151 FKGISVNAEQVRIELEPPDKAKPLGTQVQRRQKMQQKKSQDLESQVGGSYWQVRVTLI 1210
QY 373 LEEV-----SPB-LVCSMLHLCSTRLPALTVHVTPQKDGFCFCEVCKLVGYLDNRLE 424

Db	1211	LELLQHKKLRSPQILVPTLFNLLSRCL	EP-----PQEQ-----NME	1249
Qy	425	KNSTKQBIALAEEKGCSFLDPDPKQCDQFVAEYEPVLIELVEMDPSPFVCLKIGACPS	484	
Db	1250	--YTKQLILSCLLNICQL--SPDGGKIPKOILDEEKNFVELIQ-----CIRLSEMPQ	1299	
Qy	485	AKP---LLGT	492	
Db	1300	THHALLLGT	1310	

```

RESULT 15
US-11-196-400-5
; Sequence 5, Application US/11196400
; Publication No. US20050287166A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007JUSODIV
; CURRENT APPLICATION NUMBER: US/11/196,400
; CURRENT FILING DATE: 2005-08-04
; PRIOR APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,462
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: P. falciparum
US-11-196-400-5

```

	Query Match	3.7%;	Score 102;	DB 7;	Length 630;	
	Best Local Similarity	18.1%;	Frag. No. 1.1;	Mismatches 155;	Gaps 21;	
	Matches	74;	Conservative 103;			
Qy		56	PTVKSLPDCIDKDVTTAAGMLKN---	ATEEEIILYLEKT----	CDMLPKPNNSASCKE	108
		:	:	:	:	:
		:	:	:	:	:
Db		172	PTVEEI---	VAPSVESVAPSVEESVAPSVESVAENVEEIVAPSVESVAE		228
		:	:	:	:	:
		:	:	:	:	:
Qy		109	IVDSYLPLVIDIIKGEMSRPGVC	SALNLCESIQKHILAELNHOKOLESNKIPELDMTFVV		168
		:	:	:	:	:
		:	:	:	:	:
Db		229	NVEE---	SVAENVEESVAENVEESVAENVEEIVAPTVEESVATVE--EIV		282
		:	:	:	:	:
		:	:	:	:	:
Qy		169	APFM-----	ANIPLLLYPQDGRPSKPQDKONGDVQCDCIQMVTD-IQTAVRNTSTFVQAL		222
		:	:	:	:	:
		:	:	:	:	:
Db		283	APTVEESVAPTVEEIVVPVSVEESVAPSVES--	VAENVEESVAENVEESVAEN--VEESV		338
		:	:	:	:	:
		:	:	:	:	:
Qy		223	VEHVKEEC----	DRLPGMGADICKNVISYSIAIQMMHMQPKEICALVG--FCDEVK		275
		:	:	:	:	:
		:	:	:	:	:
Db		339	AENVEESVAENVEEIVAPSVESIVAPTVEE--	SVAENVATNLSDNLNLNLGGIETEEKI		396
		:	:	:	:	:
		:	:	:	:	:
Qy		276	EMPQTLVPAKVASKNWIPA-L	EVLBPBKQHEVPKASDDVYCEVFELVKLVETKLIDNNKT		334
		:	:	:	:	:
		:	:	:	:	:
Db		397	DSIINRIIEEVK--	ENVVTILEKVETTAESVTFTSNILEEQ-----	NTITNDTI	446
		:	:	:	:	:
		:	:	:	:	:
Qy		335	EKEILDAPDKCMCKLPKSL--	SEFCQEVDTYTGSSILLESVEESPELVCSMLHLCSGTR		392
		:	:	:	:	:
		:	:	:	:	:
Db		447	EKKLEELHENVLSAALENTQSEBKEVID-----	VIEEVKEEVATTIE-----		491
		:	:	:	:	:
		:	:	:	:	:
Qy		393	LPALTIVHTQPDGGCFCEVCKKL	GYGLDRNLKNS--TKQEIIAALAEK		438
		:	:	:	:	:
		:	:	:	:	:
Db		492	----TVEQAE-----	BESSTTEIFPENLEENAVESNEKVAENLEK		528
		:	:	:	:	:
		:	:	:	:	:

Search completed: January 13, 2006, 16:39:05
Job time : 12.4106 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)